

AV 서비스를 위한 코덱 기술

安 致 得
韓國電子通信研究所

I. 서론

현재 전 세계적으로 AV(Audio Visual) 서비스에 대한 연구개발이 활발히 진행되고 있는 가운데, 국내에서도 AV 서비스 개발을 위한 연구개발 사업(예, 광CATV, 고선명TV, 디지털DBS, ATM-Multi-media 단말기 개발등)이 추진되고 있다.^[1-3] AV 서비스 연구개발은 신호 정보원의 저장, 가공 및 전달을 본질로 하고 있으며, 반도체 및 컴퓨터 기술의 발전에 힘입은 디지털 신호처리 기술이 그 근간을 이루고 있다.

디지털 신호처리 기술은 AV 서비스들의 거의 전 영역에 걸쳐 미디어의 상호 통합을 용이하게 해주며, 정보 전달 방식의 공유화를 가능케하고 있다. 따라서, 향후 AV 서비스 관련 연구개발의 초점은 방식간의 상호 표준화 및 시스템의 모듈화에 맞춰지고 있는 추세이다. 방식 표준화 및 시스템 모듈화에 따른 파급 효과는 여러가지가 있을 수 있으나, 가장 중요한 것은 AV 서비스 관련 전 산업에 걸쳐 공통기반 기술을 확산시킬 수 있어 생산원가의 절감에 기여할 수 있으며, 이에 따른 관련 산업 육성 및 서비스 시장 활성화, 그리고 국제 경쟁력 강화에 있다 하겠다.

AV 서비스에 필요한 요소 기술들을 살펴보면 다음 그림 1과 같다. 그림 1에서 보는 바와 같이 AV 서비스를 달성하기 위해서는 광범위한 요소 기술들의 유기적인 결합이 필요하며, 각 AV 서비스 분야에서 공통으로 쓰일 수 있는 기술의 방식 표준화와 시스템 구현을 위한 모듈화가 중요하게 된다. 본 논문에서 다루고자 하는 AV 코덱 기술은 단말, 전송 및 저장 기술분야에 공통으로 이용되는 기술로서, AV 코덱

기술의 요체는 원신호를 저장, 가공, 전달하는 경우 신호 정보를 가능한 한 잃지 않으면서 신호가 차지하는 대역폭을 효과적으로 줄이고, 필요에 따라 다시 본래의 신호를 복원하는 것이다. AV 코덱 기술이 여타의 신호 압축 기술(예, 컴퓨터 데이터 압축 등 무손실 부/복호화)과 다른 점은 다루고자 하는 신호에 대한 인간의 시청각 특성이 고려될 수 있기 때문이다.

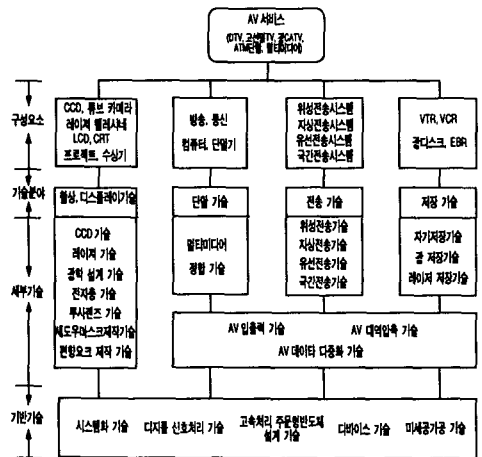


그림 1. AV 요소기술 상관도

본고에서는 AV 서비스 관련 요소 기술중 전 분야에 걸쳐 가장 공통성이 큰 부분인 오디오와 비디오의 코덱 기술에 대하여 살펴보고, 표준화 및 모듈화의 관점에서 이를 정리하고자 한다. 먼저 제 II장 및 III장에서 오디오 코덱 기술의 개요와 AV 서비스에서 현장감 재현을 위하여 사용되는 입체음향 재생 기술

에 대하여 알아본다음, 제 IV장에서 비디오 코덱 기술에 대하여 살펴보겠다. 제 V장에서는 이들 공통 요소 기술과 여러가지 AV 서비스와의 관련성과 AV 표준화에 대하여 살펴보기로 하고, 제 VI장에서는 AV서비스 개발을 위한 향후 연구개발 방향에 대하여 언급하겠다.

II. 오디오 코덱 기술

1. 개요

오디오 코덱 기술은 오디오 신호를 가공처리하여 본래의 청각 정보는 잃지 않으면서 신호가 차지하는 대역폭 또는 신호를 나타내기 위하여 필요한 데이터량을 감소시키는 신호처리 기술이다. 오디오 코덱 기술은 오디오 신호의 효율적인 기록 저장을 위해서나 또는 제한된 대역폭을 갖는 전송로로 오디오 신호를 전송하기 위하여 필요하다. 또한 고선명TV나 디지털TV 처럼 현장감을 충실히 재현하기 위하여 필요한 입체음향 구현, 음성다중, 기타 부가 오디오 서비스를 고려한 추가적인 오디오 채널이 필요한 경우 제한된 신호대역을 효과적으로 사용하기 위하여 오디오 코덱 기술이 유용하게 사용될 수 있다.

현재 사용되고 있는 디지털 음향 기기들(예, CD, DAT등)은 인간의 가청 주파수 대역(20 Hz ~ 20 kHz)에 해당하는 오디오 신호를 충분히 재생할 수 있도록 구성되어 있고, 또한 대부분의 사용자가 가청 주파수 전대역을 표현하는 오디오 신호에 익숙해져 있다. 따라서, 오디오 신호를 압축하는 경우 최소 CD 수준의 음질을 유지해야 한다. CD 수준의 음질을 유지하기 위해서는 표본화 주파수가 48 kHz(또는 44.1 kHz), 샘플당 비트수는 16 비트로서 dynamic range가 90dB 이상이 되도록 해야 한다. 이 경우 PCM을 사용하면 비트 정보량은 768 Kbit/s 가 되나, 오디오 압축 기술을 이용하면 약 128 Kbit/s 정도의 적은 비트 정보량으로 같은 효과를 얻을 수 있다. 즉, 압축된 신호를 다시 복원하여 재생했을 때 CD 수준의 음질과 거의 차이를 느낄 수 없을 정도의 음질을 유지할 수 있다.

오디오 코덱 기술이 기존의 압축 방식과 다른 가장 큰 특징은 오디오 신호 샘플간의 상관성을 이용하여 압축하는 것이 아니고, 인간의 청각특성을 이용하는 것이다. 즉, 오디오 신호를 디지털 형태로 표현하

는데 수반되는 양자화 잡음이 사람의 청각으로 지각할 수 없는 범위내에서 최대한 허용되도록 데이터 비트를 각 오디오 샘플에 가변적으로 할당함으로써 오디오 신호를 압축한다. 이때 인간 청각의 여러 중요한 특성이 이용된다.

청각의 수학적인 모델링은 일종의 multi-channel realtime analyzer 로 알려져 있다.^[4] 즉, 소리를 여러개의 가변 주파수 대역으로 분리한 다음, 각 주파수 대역마다 다른 감도를 가지고 신호를 지각한다는 것이다. 이때의 각 주파수 대역을 임계대역(critical band)이라고 하며, 이 임계대역은 저주파 대역일수록 좁아진다. 또한 청각 감도는 저주파 대역에서 고주파 대역보다 상대적으로 더 좋아진다. 오디오 압축에 이용되는 또 한가지 중요한 청각 특성은 차폐효과(masking effect)이다. 이 효과는 신호중에 레벨이 큰 신호가 발생했을 때 시간축상과 주파수축상에서의 인접한 적은 레벨의 신호들이 큰 레벨의 신호에 의해 차폐되어 들리지 않는 효과를 말한다. 실험적으로 시간축상의 경우에 큰 레벨의 신호 발생 시점에서 이전 시간으로는 약 4ms, 이후 시간으로는 약 40 ms의 차폐효과가 있고, 주파수축 상의 경우는 큰 레벨의 신호 발생 주파수에서 저주파 대역으로의 차폐 범위보다 고주파 대역으로의 차폐 범위가 더 크고, 전체적인 차폐 범위의 크기는 저주파 대역의 신호보다 고주파수 대역의 신호에서 더 크다.^[4] 이러한 실험 결과는 청각 특성에서 각 임계대역 별로 각각 다른 감도를 가지고 소리를 지각한다는 청각 모델링과 일치하며, 저주파 대역에서 더 민감하다는 청각 특성을 반영하고 있다.

2. 오디오 코덱 예

대부분의 오디오 코덱은 인간의 청각 특성을 이용하여 압축을 하며, 이러한 코딩방식을 perceptual coding이라 한다. Perceptual coding 방식에는 크게 두가지가 있는데, 하나는 대역분할 부호화(subband coding) 방식이며, 다른 하나는 변환 부호화(transform coding) 방식이다. 유럽의 EUREKA 147 프로젝트 결과로 나온 MUSICAM(Masking pattern Universal Subband Integrated Coding And Multiplexing)은 대역분할 부호화 방식을 사용하는 perceptual coding 방식의 대표적인 예이다.^[5] MUSICAM 방식은 원래 DAB(Digital Audio Broadcasting)에서 사용할 목적으로 개발되었는데,

현재는 MPEG 오디오 방식과 DCC(Digital Compact Cassette) 방식 등에 이용되고 있다. 한편 Dolby에서 개발한 AC-2^[4]와 미국, 프랑스, 독일에서 공동개발한 ASPEC^[6]은 transform coding 방식을 사용하는 대표적인 예이다. 그 중 ASPEC은 현재 MPEG-1 오디오의 계층 III 압축방식에 사용되고 있다.^[9]

MUSICAM 방식의 기본 구조는 그림 2와 같다.^[7] 예로서, 표본화 주파수가 48 kHz 인 오디오 인코더의 경우, 먼저 입력된 오디오 샘플에 대해서 시간축상에서 8 ms 에 해당하는 시간 프레임으로 나누고, 이를 대역 분석 필터를 이용하여 주파수축상에서 동일 대역폭을 갖는 32개의 서브밴드로 나눈다. 각 서브밴드(48kHz 샘플링의 경우 12개의 샘플로 구성)에 대하여 분석된 신호 특성과 FFT로부터 얻어진 신호 특성을 이용하여 차폐 문턱값(masking threshold)을 구하고, 이를 근거로 하여 각 서브밴드의 샘플들에 대해서 양자화 비트를 할당한 다음 정규화를 이용한 블록 압신(block-companding) 방식으로 부호화한다. 디코더의 경우는 비트 할당 정보를 이용하여 부호화된 샘플을 복원하고, 이를 대역 합성 필터를 통해 시간축상의 오디오 샘플로 변환시킨다.

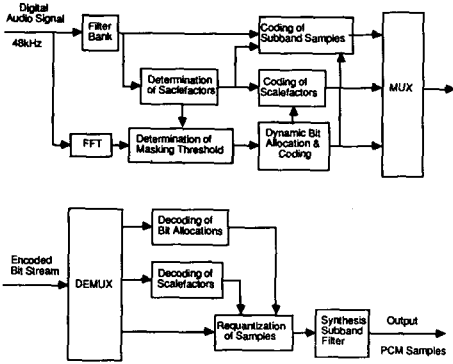


그림 2. MUSICAM 인코더와 디코더의 기본 구조

MUSICAM 방식은 대역 분할 필터를 통해 동일 대역폭을 갖는 서브밴드로 나누어서 처리하기 때문에, 특히 저주파에서 서브밴드보다 좁은 대역폭을 갖는 임계대역들에 대해서 정확한 분석이 어렵다. 따라서 이를 보상하기 위해 FFT를 사용하게 되는데, 이로 인해 인코더가 복잡해지는 단점이 있다. AC-2 방식은 그림 3과 같은 기본 구조를 갖는다.

^[4] 입력된 오디오 샘플에 대해서 overlapped window와 MDCT(Modified Discrete Cosine Transform) 및 MDST(Modified Discrete Sine Transform)로 구성되는 TDAC(Time Domain Aliasing Cancellation)^[8]를 이용하여 주파수 변환을 수행한 다음, 인접하는 주파수 계수들을 모아 각 임계대역에 해당하는 서로 다른 대역폭의 서브밴드를 구성한다. 각 서브밴드 내의 주파수 계수들은 하나의 대표값에 해당하는 지수(exponent)와 여러개의 소수(mantissa)를 갖는 frequency block floating-point 표현으로 변환된 다음, 전체 서브밴드의 지수들을 이용하여 오디오 샘플들에 대한 log-spectral envelope를 추정한다. 이 envelope 정보로부터 각 서브밴드에 대한 청각상의 중요도를 추정하고, 이에 따라 서로 다른 비트를 할당하여 해당 소수들에 대해 양자화를 수행한다.

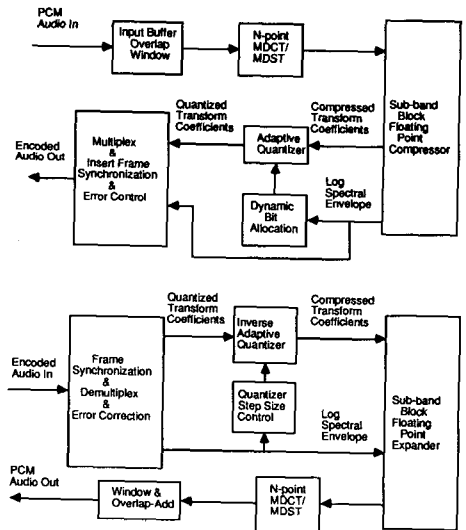


그림 3. AC-2 인코더와 디코더의 기본 구조

오디오 압축 방식이 실용화되기 위해서는 압축 성능뿐만 아니라 실시간의 구현성, 적은 지연 시간, 오차에 대한 강인성, 간단한 시스템의 구현(특히 디코더) 등이 고려되어야 한다. 앞에서 설명한 두가지 압축 방식은 대체로 이러한 내용들을 만족시키는 것으로 알려져 있다. 표 1에서 이들 두가지 방식의 특성을 비교하였다.

표 1. 오디오 압축 방식 비교

항목 \ 방식	MUSICAM	AC-2
압축방식	Perceptual Coding (Subband Coding)	Perceptual Coding (Transform Coding)
Filter bank 구성방법	Polyphase Network 이용	Evenly-stacked TDAC 이용
Subband	Equal bandwidth	Nonequal bandwidth
Subband 당 샘플수	12 샘플	임계대역 대응 가변
시간프레임	8 ms (384 샘플)	10.6 ms (512 샘플)
Secondary Filter bank	필요(FFT)	불필요
Bit 할당 방법	Mask-to-Noise Ratio 이용	Log Spectral Envelope 이용
Masking Modelling	복잡	간단
부호화	Saclefactors Normalized subband samples	Exponents Mantissas
Complexity	복잡	간단

III. 입체음향 재생 기술

입체음향에 관한 연구는 1930년대 초 영국의 Blumlein과 미국의 Snow에 의해 시작되었으며, 상용적인 구현은 1950년대 미국에서 관객을 TV에 뺏기는 것을 막기 위한 방편으로 영화업자들에 의해 본격적으로 도입이 시작되었다. [10] 입체음향에는 모노포닉(mono phonic)에 비해 크게 두가지 중요한 효과, 즉 입체 음상 (stereophonic imaging) 효과와 청각적 효과(auditory impression)가 있다. [11]

입체 음상 효과란 오케스트라의 연주를 스테레오로 청취할 때 각 악기의 음상을 정확히 지각할 수 있는 것과 같이 여러개의 음원(sound source)의 위치를 독립적으로 분리하여 청각적으로 지각할 수 있도록 하는 음상(sound image)을 제공하는 효과를 말한다. 청각적 효과란 입체 음상 효과에 비하여 복잡한 효과로서 크게 공간적 효과(spatial impression)와 효과음(effect sound)으로 구분된다. 공간적 효과에는 여러 방향으로 부터의 반사음을 재생함으로써 마치 콘서트 홀에 있는 것과 같은 느낌을 주는 방법과 잔향음을 이용하여 소리의 확산효과를 제공함으로써 실제로 존재하는 공간보다 더 큰 공간에서 음을 듣는 듯한 느낌을 주는 방법이 있다. 전자는 초기 반사음(early reflections)에 의해, 그리고 후자는 지연 시간이 긴 잔향음(reverberation)에 의해 효과를 얻을 수 있다. [11] 한편, 효과음은 콘서트 홀에서 처럼 박

수 소리나 웃음 소리, 빗 소리, 바람 소리 등을 현장감있게 재현하는 것을 말한다.

입체음향의 효과적인 재생은 여러 음원들로 부터 구성되는 음장 (sound field)을 어떻게 충실하게 재현하는가에 달려 있다. 입체음향의 실제적인 구현은 다수의 스피커를 이용한 음의 재생을 통해 이루어지는데 입체 음상은 전면 스피커로 재생되고, 청각적 효과는 서라운드 스피커에 의해 재생된다. 기존 두채널 스테레오의 가장 큰 문제점은 방향감이 불안정(directional unstability) 하다는 점이다. 즉 두채널 스테레오에서는 두 스피커 사이의 중앙 위치에서 듣는 경우는 문제가 없지만, 중앙에서 조금만 벗어나도 먼저 도달하는 스피커의 소리에 의해 크게 방향성감지에 영향을 받는 선행 효과(precedence effect)에 의해 음상 정위가 불안정해진다. 그동안 두채널 스테레오가 무리없이 사용되어 왔던 이유는 입체 음상 효과보다는 청각적 효과에 어느 정도 만족할 수 있었기 때문이나, 현장감을 재현하기 위해 필요한 영상과 소리의 일치성(directional coincidence)의 구현에 있어서 방향성의 불안정은 특히 문제가 된다.

두채널 스테레오 재생이 가지고 있는 음상의 불안정함을 해결하기 위한 여러가지 실험과 방법들이 연구되어 왔다. [11-14] 이들 실험은 전면 스피커의 갯수와 위치를 변화시키면서 방향감의 오차를 측정하는 방법인데, 두채널에서 세채널(좌, 우, 중앙)로 증가시키면 방향감이 안정되는 효과가 얻어지지만, 네채널로 증가시킬 경우는 그다지 큰 효과가 없다는 점이 실험의 공통된 결과이다. 이를 종합해 보면 스피커의 수가 많을수록 입체 음상을 재현하는데 유리하지만 허용되는 방향성의 오차, 가격대 성능비, 전송 비용 등을 고려하면 대체로 세개의 전면 스피커를 사용하는 것이 적당하다.

두채널 스테레오 재생은 청각적 효과를 충분히 재현하는데 있어서도 단점이 있다. 청각적 효과를 충실하게 재현하기 위해서는 서라운드 스피커를 이용하게 되는데, 이에 관한 실험도 역시 서라운드 스피커의 갯수와 위치를 가변시키면서 청각적 효과를 측정한다. [11-12] 이러한 실험들을 종합할 때 전면 스피커와 마찬가지로 가정에서의 사용을 전제로 적당한 수준의 청각적 효과, 가격대 성능비, 전송 비용 등을 고려하면 한쌍의 서라운드 스피커의 사용이 적당하다.

전면 스피커와 서라운드 스피커 이외에도 저주파수 대역(200 Hz 이하)의 오디오 신호를 효과적으로 재

생하기 위한 서브우퍼(subwoofer) 스피커의 사용도 고려되고 있다. 저주파수 대역에서는 청각의 특성상 상대적으로 높은 신호 레벨을 필요로 하며, 또한 무지향성이므로 전면에 별도로 한개의 서브우퍼 스피커를 배치하는 것이 효과적이다.^[10] 이상을 종합하면 효과적으로 현장감을 구현하기 위한 오디오 채널의 구성을 그림 4와 같이 하는 것이 바람직하다.

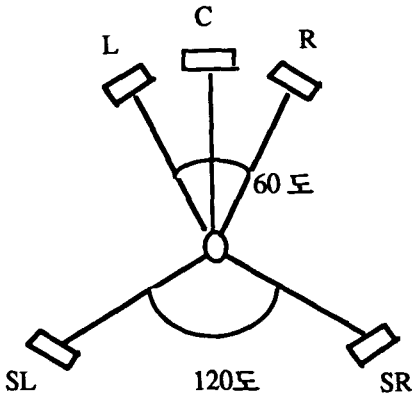


그림 4. 고선명 TV 오디오 채널 구성

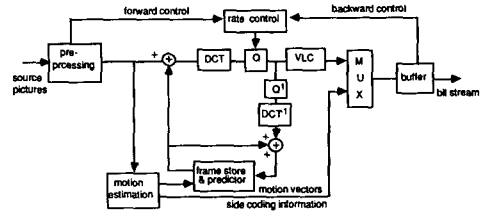
현장감을 충실하게 재현하기 위한 최적의 오디오 채널 구성과 함께 입체 음향의 재생에 있어서, 또한 가지 고려할 점은 서로 다른 채널 구성을 갖는 시스템간의 채널 호환성을 들 수 있다. 채널 호환에는 적은 채널 수에서 많은 채널 수로 채널 구성을 변환하는 상향 호환(upward compatibility)과 많은 채널 수에서 적은 채널 수로 채널 구성을 변환하는 하향 호환(downward compatibility) 방식이 있다. 호환성에 대해서는 여러가지 실험 결과들이 발표되고 있지만, 경우에 따라 성능이 변하기 때문에 추후 더 많은 실험이 필요한 분야이다.

IV. 비디오 코덱 기술

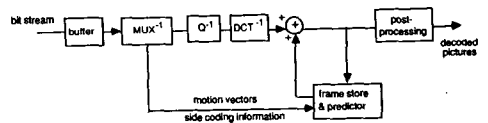
1. 개요

일반적으로 영상 신호는 수십 Mbps에서 고선명 TV의 경우 1 Gbps 이상의 방대한 데이터량을 갖는다. 따라서 기존 전송로에서의 전송을 위해서나, 제한된 대역폭을 갖는 저장장치에 기록하기 위하여는 수십:1 이상의 데이터 압축을 요하며, 설사 충분한 전송로나 저장장치가 제공된다 하더라도 많은 중복성

을 갖는 영상신호를 그대로 전송/기록하는 것은 비효율적이다. 영상 데이터 압축은 기본적으로 영상 신호가 포함하는 많은 중복성을 제거함으로써 이루어지며, 이 경우 데이터의 손실 없이 압축하는 것을 무손실 부호화(lossless coding)라고 하고, 인간의 시각 특성을 함께 고려하여 인지도가 떨어지는 부분을 중심으로 손실을 감수하여 더 높은 압축율을 얻는 것을 손실 부호화(lossy coding)라 한다. 이들 중복성을 제거하는 압축 기법은 시간적 중복성을 제거하는 방법과 공간적 중복성을 제거하는 방법으로 크게 구분된다.



(a) 부호화기



(b) 복호화기

그림 5. DPCM/DCT 복합 부/복호화기 기본구조

영상 부호화 방식에는 대역분할 부호화(sub-band coding), 벡터 양자화(vector quantization)에 근거한 부호화, 멀티미디어 분야의 대표격인 CD-I, DVI의 RTV(Real Time Video), PLV(Presentation Level Video)등 다른 여러 부호화 기법이 있지만 현재 주류를 이루는 방식은 움직임 보상(motion compensation) 예측 기법과 DCT(Discrete Cosine Transform)를 함께 사용하는 DPCM/DCT 복합 부호화 방식이다.(그림 5 참조) CCITT의 H. 261^[15] 및 ISO/IEC의 MPEG^[9] [16] 에서 DPCM/DCT 복합 부호화 기법을 사용하고 있으며, 미국 HDTV 표준 제정을 위하여 제안된 4가지 디지털 방식들이 모두 이방식을 근거로 하고 있다.^[17-19] 또한 이 부호화 기법을 실시간 구현하기위한 반도체 기술도 상당

히 진척되어 실제 상업용 칩세트들도 나오고 있다. 이러한 현재의 추세로 볼때 당분간은 이들 기법을 기본으로 하는 동영상 압축기법이 주류를 이룰 것으로 보인다. 본고에서는 이동 보상 DPCM/DCT를 중심으로 하여 비디오 코덱 기술을 전/후처리, 데이터 압축/복원, 버퍼제어 부분으로 나누어 살펴 보겠다.

2. 전/후처리

통상 비디오 코덱에서는 전처리(pre-processing) 과정을 통하여 전체 코덱의 부호화 성능을 높일 뿐만 아니라, 시-공간 영역의 부표본화(sub-sampling)를 통하여 데이터 압축을 행한다. 먼저 중복성이 많은 R, G, B 입력신호를 시각 특성에 맞추어 휘도신호(Y)와 색신호(Cb, Cr)의 색체계로 변환한다. 이 경우 정보의 큰 손실 없이 색신호(Cb, Cr)의 샘플수를 줄여 많은 데이터 감축을 하는데 보통 방송 서비스등을 위한 4:2:2 또는 통신 서비스등을 위한 4:2:0의 영상양식을 많이 사용한다. 색신호에 대한 감축 뿐만 아니라 전처리 과정에서 시-공간의 부표본화를 통하여 높은 데이터 압축율을 얻기도 하는데 저 전송율 코덱에서 특히 유용하다. 또한 데이터 압축부와 연계하여 양자화 간격등의 부호화 정보에 따라 적응적으로 저역통과 필터링의 전처리를 하거나 신호원에 포함된 잡음제거를 통하여 코덱의 부호화 효율을 향상시킬 수도 있다

후처리(post-processing) 과정은 비디오 코덱의 복호화기를 통하여 재생된 영상의 화질을 향상시키기 위하여 부호화 과정에서 생긴 blocking artifact등의 부호화 잡음을 제거하거나, 전처리 과정에서 부표본화를 한 경우 디스플레이를 위한 영상양식으로 보간하는 등 영상 감축과는 별도의 과정이나, 비디오 코덱의 전체 성능에 미치는 영향이 매우 크다.

3. 데이터 압축/복원

1) 움직임 보상 예측 부호화

프레임간(inter-frame) 예측 부호화는 시간상의 중복성을 제거하여 데이터를 압축하는 방법으로, 특히 물체의 움직임에 기인한 변화를 보상하여 예측하는 움직임 보상 DPCM은 많은 데이터량을 감축할 수 있다. 움직임 보상 부호화에서는 정확한 움직임을 검출하여 예측 오차를 줄이는 것이 중요하며, 움직임 검출 방법은 블럭정합 알고리즘(BMA:Block Matching Algorithm)과 화소순환 알고리즘(PRA:Pixel

Recursive Algorithm)으로 나눈다. BMA는 한 영상을 여러개의 작은 블럭으로 나누어 각 블럭 단위로 움직임을 추정하는 것이고, PRA는 각 화소 단위로 움직임을 추정하는 것으로 BMA 보다 비교적 정확한 움직임을 검출하지만 시스템의 복잡도와 계산량이 크기 때문에 현재는 BMA에 의한 움직임 검출이 대부분의 비디오 코덱에서 사용되고 있다.

BMA에서 탐색 블럭의 크기, 탐색 방법 및 탐색 영역의 범위에 따라 움직임 검출의 정확도와 계산량이 상반된 관계를 갖는다. 블럭의 크기가 작을 수록 움직임 예측의 성능이 증가하지만 각 블럭 마다 움직임 벡터를 전송해야하므로 움직임 벡터 정보의 증가로 인하여 부호화 효율이 떨어질 수 있다. 일반적으로 BMA에서는 8X8 이나 16x16 크기의 블럭을 사용한다. 한편, 전영역 탐색(full search)이 탐색 영역내에서 움직임 벡터를 가장 정확하게 찾아내지만 많은 계산량을 요하기 때문에 TSS(Three Step Search)등의 고속 탐색 기법들이 연구되어 왔다. 그러나 VLSI 칩 설계에서 병렬처리 기법이 발달함에 따라 병렬 처리에 용이한 전영역 탐색 기법이 MPEG등의 여러코덱에서 사용되고 있다.

MPEG-2와 ATSC의 ADTV 방식에서는 정수화소 단위의 정밀도를 갖는 움직임 벡터를 추정후 반화소 단위의 움직임 벡터 까지 추정하여 정확도를 높이고, 격행주사된 영상의 경우 프레임간의 움직임 추정 뿐만아니라 필드간의 움직임도 고려하는 dual field prediction, FAMC(field time adjusted MC)등 다양한 추정기법들이 제안되었다. 또한 B 프레임 도입하여 양방향(bi-directional prediction)으로 움직임을 추정함으로써 uncovered area에서도 적절한 움직임 추정이 가능하게 되어 예측 오차를 더욱 줄이고 있다. 결론적으로 움직임 추정의 정확도는 연산의 복잡도와 직결되므로 그 만큼 하드웨어 구현 기술이 중요한 요소이며, 하드웨어 구현을 위한 시스템 차원의 계속적인 연구가 요구된다.

2) 변환부호화 및 양자화

변환을 이용한 데이터 압축 기법은 원영상에 분산된 에너지를 몇 개의 변환 계수에 집중시켜 공간상의 중복성을 제거하는 기법으로, 공분산에 근거한 KLT(Karhunen Loeve Transform)가 에너지 집중과 비상관성 측면에서 이론상 최적의 변환 기법이지만 구현이 어려우므로 실제의 영상 부호화에서는 KLT와 가장 근접한 성능을 갖는 DCT가 많이 사용되고 있

다. DCT의 블럭의 크기는 에너지의 집중과 복잡도의 고려에서 8x8이 가장 효율적인 것으로 알려져 있다.^[20]

블럭 단위로 변환을 하는 DCT와는 달리 대역분할 부호화(sub-band coding)는 QMF(Quadrature Mirror Filter)를 이용하여 한 영상을 몇개 대역의 영상으로 나눈후, 시각 특성을 이용하여 각 분할된 영상을 독립적으로 부호화 하는 것으로, 최근에는 계층적 채널 구조를 갖는 BISDN에서 효과적인 부호화 기법으로 많이 연구되고 있다.^[21] 대역 분할 부호화와 유사한 것으로 wavelet 변환을 이용하여 좀더 효과적인 영상 분할을 피하는 부호화 기법^[22]과 blocking artifact 감소를 위한 방법으로 LOT(Lapped Orthogonal Transform)^[23]가 연구되고 있다. 이동 보상 DPCM 방식에서는 움직임 추정이 블럭 단위로 이루어지므로 거의 모든 비디오 코덱이 DCT를 사용하여 블럭단위의 예측 오차를 변환하여 시-공간 중복성을 제거하고 있다.

예측 오차의 변환 계수들은 양자화 단계를 거쳐 전송하게되는데 DPCM/DCT 복합부호화의 부호화 오차는 양자화 과정에서 발생한다. 변환계수의 통계적 특성이 입력 영상에 따라 일정하지 않아 최적의 양자화를 설계하기가 어렵고, 단순히 통계적 특성에 따라 양자화 잡음을 최소화하는 양자화를 설계하기 보다는 시각 특성을 고려하여 시각적으로 양자화 왜곡을 최소화하는 방법을 많이 사용하고 있다. 이를 위해서 DCT계수들은 먼저 인간의 시각 특성에 따른 가중행렬(weighting matrix)에 의해 나뉘어진후 양자화된다.

3) 엔트로피 부호화

움직임 벡터, DCT 계수 및 기타 부가정보등의 부호화 정보들은 복호기와 미리 약속된 부호를 이용하여 전송한다. 이때 전송하여야할 비트 발생을 최소화 하기위하여 엔트로피 부호화를 한다. 엔트로피 부호화는 입력 영상의 통계적 특성을 이용하는 것으로 양자화 과정과는 달리 무손실 부호화 알고리즘으로 가변장 부호화(VLC : Variable Length Coding) 기법과 줄길이 부호화(RLC : Run Length Coding) 기법이 널리 사용되며, H.261이나 MPEG, JPEG에서도 가변장 부호화 및 줄길이 부호화를 함께 사용한다.

가변장 부호로는 허프만 부호가 대표적이며, MPEG-2의 경우 MBA(Macro Block Address increment),

MVD(Motion Vector Difference), MTYPE (Macroblock Type)등의 부호화 정보를 허프만 부호를 이용하여 부호화 한다. 양자화된 DCT 계수는 지그재그 주사를 하여 2차원 줄길이 부호화 기법의 (run, level) 쌍으로 표현한 다음 가변장 부호화를 한다. 한편, 통계적 빈도가 매우 낮은 데이터는 고정장 부호화(fixed length coding)를 한다. 엔트로피 부호화된 정보들은 정해진 문법(syntax)에 따라 비디오 비트열(bit stream)을 형성하여 출력된다.

4. 비트 발생률 제어

부호화 정보들을 엔트로피 부호화하면 발생하는 데이터량은 데이터의 통계적 특성에 따라 매우 가변적이 되므로 고정된 대역의 채널로 출력시키기 위해서는 발생 비트율의 완충 역할을 위한 버퍼가 필요하다. 이때 버퍼의 넘침(overflow)이나 결핍(underflow)이 발생하지 않도록 비트 발생을 제어하여야 한다. 비트 발생 제어는 주로 버퍼 상태에 따라 양자화의 간격을 조절함으로써 이루어지는데, 양자화 간격의 변화는 화질에 직접 영향을 미치므로 화질의 저하를 최소화 하면서도 일정한 화질을 유지하도록 버퍼 제어를 하여야 한다. 이를 위하여 버퍼로 부터 버퍼의 충만도에 따른 피드백(feedback) 제어 뿐만아니라, 입력 영상의 통계적 특성에 따른 인간의 시각 특성을 고려하여 적응적으로 비트 발생율을 제어하기위한 전처리 과정으로 부터의 전방향 제어(forward control)를 함께 사용하는 경우가 많다.

V. AV 서비스 호환성

1. AV 서비스 통합

AV 서비스 관련 기술의 디지털화는 미디어의 상호 통합을 용이하게 해주고 있다. 가장 대표적인 예는 최근의 컴퓨터를 기본으로한 멀티미디어 서비스이다.^[24] 대부분의 AV 서비스에서 보듯이 디지털 기술의 도입은 제작(production) 측에서 먼저 시작되어 일반 소비자측으로 전이하는 과정을 밟는다. 예를들어, CD, DAT, TV프로그램 제작등이 그러하다. 이것은 제작측의 수요가 작아서 시장에 대한 영향이 상대적으로 적기 때문이다. 그러나, 제작측의 초기 디지털 기술 도입은 비용이 많이 드는 단점이 있다. 통신에서의 AV 서비스의 경우는 망과 교환기의 디지털화와

직접 연관이 있다. 현재 외국에서 시험운영되고 있는 ISDN용 AV 서비스나 향후의 B-ISDN용 AV 서비스가 대표적인 예이다.^[25] 또한 케이블 TV 서비스와 정지위성을 통한 디지털 TV 방송, 위성 또는 지상파를 이용한 디지털 오디오 방송(DAB : Digital Audio Broadcasting)등이 실용화 단계에 있거나, 서비스가 시작되고 있다.^[26] 이밖에 CD, VTR 에 의한 AV서비스나 영화, 사진, 인쇄, 출판등에서도 디지털화가 이루어지고 있다.

AV 서비스의 영향력에서 볼때 가장 큰 파급효과를 미치는 서비스는 방송과 통신 서비스이다. 방송은 스튜디오 장비의 디지털화와 함께 전송의 디지털화가 시작되었으며, 고속, 광대역전송에 의한 다양한 서비스들(예, 비디오 쇼핑, 문자뉴스, 오디오 그래픽등)의 경제적인 제공이 가능해지고 있다. 통신은 망의 디지털화에 이어 단말의 디지털화에 따른 고품질화 및 사용자의 부가 서비스에 대한 욕구가 증대되고 있다. 방송과 통신 공히 디지털 AV 서비스 관련 기술의 발전과 사용자 요구의 다양화에 힘입어 상호 밀접하게 결합하고 있는 추세이다. 이에따라 AV 공통기술의 상호 표준화에 따라 핵심기술의 공유가 가능해지고 있으며, 통신망을 통한 방송서비스 제공이 가능해짐으로써, 다양하고, 경제적인 AV 서비스를 제공할 수 있게 되었다.

육성을 촉진하여 국제 경쟁력을 제고할 수 있게 해준다. 따라서, 기술적인 고려 이외에 방송과 통신의 상호 발전 및 통합에 따른 방송/통신 환경의 변화에 능동적으로 대처할 수 있는 표준화 및 제도적 뒷받침이 이루어져야 할 것이다.^[27]

2. AV 서비스 표준화

AV 서비스에서 디지털 기술의 도입이 새로운 서비스 창출이나 기존 서비스의 품질을 높일수 있다는 것은 주지의 사실이다. 디지털화는 경우에따라 기존 서비스 또는 새로운 디지털 AV 서비스와 호환성의 문제를 야기시킬 수 있다. 호환성의 문제를 해결하기 위하여는 앞절에서 언급한 바와 같이 최대한 공통의 방식 표준화를 추구해야 한다. 통신 서비스의 경우에는 국제 표준화에 대한 활동이 활발하게 이루어져 왔으며, 이의 대표적인 예가 CCITT H.261^[15] 표준과, 현재 진행중인 H.26X 표준이다.^[16] 특히 상기의 표준들은 ISO/IEC와 긴밀한 협조하에 추진되고 있으며, 여타의 디지털 AV 서비스 도입을 위하여 좋은 보기가 되고 있다. 방송/통신의 입장에서 보면 AV 서비스 방식표준화를 그림 7과 같이 나타낼 수 있으며, 특히 현재 국내에서 활발하게 연구개발되고 있는 DTV(Digital TV), 고선명TV, 광CATV 서비스가 대표적인 예이다. 그림 8은 AV 서비스의 발전 전망과 이와 관련된 국제표준을 도시한 것이다.

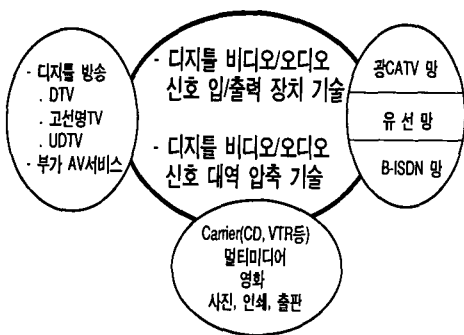


그림 6. AV 서비스의 공통기술

그림 6에서 보는 바와같이 AV 서비스 통합을 가능케 해주는 것은 바로 공통의 AV 신호처리 기술, 즉, 디지털 오디오/비디오 입출력장치 기술 및 신호압축 기술이다. 서론에서도 언급했듯이 상호 공통기술의 추구는 서비스 보급 활성화 뿐만 아니라 관련 산업의

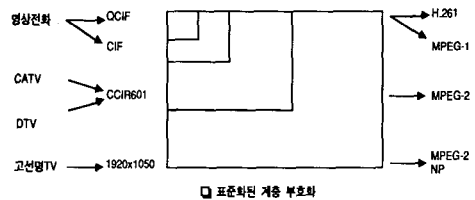
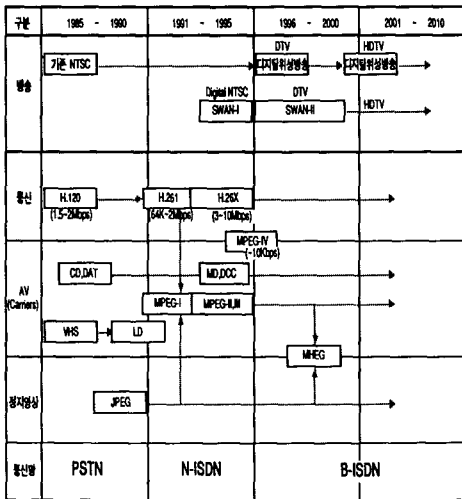


그림 7. AV 표준화 개요

방식 표준화가 AV 서비스의 상호 호환성을 해결해 줄 수 있는 것 이외에 표준화는 또한 AV 서비스의 모듈화된 구현을 가능하게 해준다. 서비스 구현의 모듈화는 특히 관련 산업의 파급효과에 의한 저가격화 및 기술 확산에 도움을 준다. 이상을 종합하면 AV 서비스 개발을 위한 바람직한 일반적인 접근 방법은 다음과 같다.

- 매체별 방식의 공통 부분을 최대한 표준화
 - 국제 표준의 적극 수용
- AV 서비스의 경우 특히 다음의 기술적 사항들을 중요하게 고려해야 한다. 즉,
- 서로다른 매체에 공통으로 사용될 수 있는 AV 압축 방식과 신호 및 프로그램 다중화 방안
 - 다양한 서비스에 맞는 융통성 있는 시스템 구조 등이 충분히 검토되어야 한다.



JPEG: Joint Photographic coding Experts Group
 MPEG: Moving Picture coding Experts Group
 MHEG: Multimedia Hypertext information coding Experts Group
 CD: Compact Disk
 DAT: Digital Audio Tape
 MD: Mini Disk
 DCC: Digital Compact Cassette

그림 8. AV 서비스 향후 전망

VI. 향후 연구개발 방향

서두에서 언급했듯이 현재 국내에서는 차세대 AV 서비스 도입을 위한 활발한 연구 개발이 이루어지고 있다. 먼저 국내 방송 및 통신 AV 서비스의 획기적인 전환점이 될 DTV 위성 방송 서비스와 광CATV 서비스를 들 수 있다. 우리나라 최초의 상용위성이 될 무궁화 위성이 1995년 발사되어 96년부터 본격적인 서비스에 들어갈 예정으로 있으며, 이와 관련하여 정부는 최근에 위성방송 전송방식을 디지털 방식으로 하기로 결정하였다. DTV 위성방송 서비스를 위한

관련 시스템의 연구개발은 한국통신, 방송사, 사업체 및 ETRI가 공동으로 수행중이다. 한편, 광CATV 시스템은 ETRI와 산업체를 중심으로 현재 실험시제품 개발이 완료된 상태이며, '95년부터 시범 서비스에 들어갈 예정이다. 이밖에 차세대 AV 서비스로서 고선명TV의 방송/통신 서비스에 대한 연구개발도 활발히 이루어지고 있다.^[2-3] 특히, 최근 생산 기술원과 산업체가 주축이 되어 고선명TV 수상기 시제품을 개발하고 대전 EXPO에 전시중이다. 한편, 고선명TV 방송 전송시스템에 대한 연구개발도 DTV 위성방송 전송시스템과 마찬가지로 국내 관련 연구소와 산업체가 주축이 되어 수행되고 있다.

이러한 AV 서비스를 위한 시스템 연구개발에 있어서 중요한 것은 제V장에서 언급한 바와 같이 각종 AV 서비스 방식들 사이의 호환성 확보라 할수 있다. 이 중에서 특히 문제가 되는 것은 수신단에서의 디코더 및 디스플레이 장치의 호환성이다. 새삼 강조할 필요없이 AV 서비스 방식간의 호환성 확보 없이는 서비스 모듈화는 물론 상이한 서비스들의 확산과 관련 산업의 균형적인 발전에 막대한 지장을 초래할 것이다. 따라서, 향후 AV 서비스 관련 연구 개발시 고려해야 할 점은 여러가지가 있으나^[28], 가장 중요한 것은 새로운 서비스와 또한 이후에 도래할 서비스들이 가장 최소의 비용으로 수용될 수 있는 방향으로 연구개발이 이루어져야 한다는 것이다.

그림 9는 다양한 AV서비스들 사이의 호환성 확보 시 예상할 수 있는 AV서비스 운용 개념도이다. 주목할 것은 사용자에게 필수적인 각종 AV 서비스를 위한 디코더와 디스플레이 장치의 호환성 확보이다. 특히 방송 방식은 한번 결정되면 최소 30~40년 이상 사용되므로 특별한 주의를 요한다. 비디오 코덱의 경우 호환성 확보를 위하여 전세계적으로 여러가지 방식이 연구되고 있으나, 가장 바람직한 접근 방법은 MPEG-2에서 고려되고 있는 계층적 부호화 방식이다.^[16-9] 오디오 코덱의 경우는 비디오 코덱보다 상대적으로 낮은 비중을 차지하나, 상이한 AV 서비스들 사이의 호환성 측면에서는 예외가 될 수 없다. 현재 전자통신연구소에서는 방송사, 한국통신, 산업체와 함께 상기의 호환성 확보를 고려하여 DTV/고선명TV/광CATV 시스템에 공통으로 사용될 수 있는 AV 용 코덱 방식에 대한 연구개발이 이루어지고 있으며, 차후 이를 기반으로 하여 향후 AV 서비스에 필요한 시스템을 개발할 예정이다.

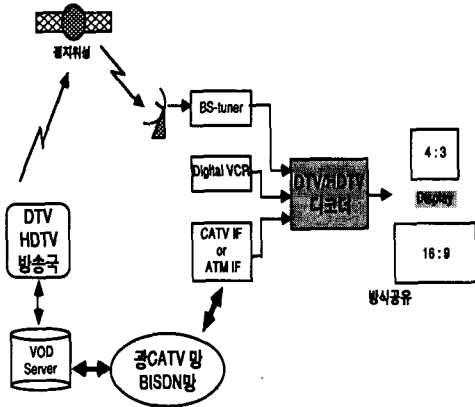


그림 9. 차세대 AV 서비스 방식 공유

Ⅶ. 결어

본고에서는 AV 서비스를 위한 오디오, 비디오 코덱 기술에 대한 개요와 AV 서비스에서 중요한 현장감을 충실히 재현하기 위한 입체음향재생 기술에 대하여 살펴보았다. 디지털화에 따른 AV 서비스 발전은 산업, 경제적 측면에서 뿐만 아니라 문화, 사회적으로도 지대한 영향을 끼칠 것으로 예상되고 있으므로, 앞으로 지속적인 관심과 연구 개발이 필요하다. 향후 순조로운 AV 서비스 확산과 관련 산업의 활성화를 위해서는 선진 제국의 동향을 파악하고, 국제표준화에 적극 참여하여 산업의 국제 경쟁력을 강화할 수 있는 방향으로 노력하여야 한다. 특히 AV 서비스 관련 기술의 발전에 따른 AV 서비스 통합과 표준화에 대하여 많은 관심을 기울여야 하며, 종합적인 연구 개발과 함께 사회제도적인 뒷받침이 동시에 이루어져야 할 것이다.

감사의 글

본 논문을 작성하기 위하여 도움을 주신 강철희 박사님과 영상통신연구실원 여러분에게 감사드립니다.

參考文獻

[1] 광CATV 시스템 기술개발, 1992 연차보고서,

한국전자통신연구소, 1992.
 [2] 고선명TV 전송 기술개발, 1992 연차보고서, 한국전자통신연구소, 1992.
 [3] 광대역 종합 정보통신망 연구개발사업, 한국통신, 1992.
 [4] L. Fielder, G. Davidson, "AC-2 : A Family of Low-Complexity Transform-based Music Coders", 1991 AES Workshop on Digital Audio, 1991.
 [5] Yves-Francois Derierey, "MUSICAM Source Coding", Proceedings of the 10th International AES Conference, pp. 71-79, Sep. 1991.
 [6] K. Brandenburg, "ASPEC Coding", Proceedings of the 10th International AES Conference, pp. 81-90, Sep. 1991.
 [7] G. Stoll, Y.F. Dehery, "High-quality Audio Bit-rate Reduction System Family for Different Applications", Proc. of IEEE Intl. Conf. on Comm., pp. 937-941, 1990.
 [8] J.P. Prince, A.B. Bradley, "Analysis/Synthesis Filter Bank Design Based on Time-domain Aliasing Cancellation", IEEE Trans. on Acous., Speech, and Sig. Proc., vol. ASSP-34, no. 5, pp. 1153-1161, Oct. 1986.
 [9] ISO/IEC, MPEG-1 DIS 11172, 1992.
 [10] Tomlinson Holman, "Psychoacoustics of Multi-channel Sound Systems for Advanced Television", 1992 NAB HDTV World Conference Proceedings, pp. 13-20, 1992.
 [11] Gunther Thelle, "HDTV Sound Systems : How Many Channels?", AES 9th International Conference, pp. 217-232, 1991.
 [12] D. G. Kirby, "Experiences with Multi-channel Sound for HDTV", Proceedings of IBC '92, pp. 528-532, 1992.
 [13] Graham Carter, "Multi-channel Sound for HDTV", AES 10th International Conference, pp. 135-139, 1992.

- [14] K. Nakabayashi, K. Kurozumi, E. Miyasaka, S. Komiyama, A. Morita, "Three-one Quadraphonic Sound System for High Definition Television", AES 10th International Conference, pp. 141-146, 1992.
- [15] Video Codec for Audio Visual Services at p x 64 Kbits/s, CCITT Recomm. H. 261, 1990.
- [16] Generic Coding of Moving Pictures and Associated Audio, ISO/IEC 11172-6, May 1993.
- [17] Special Issue on HDTV broadcasting, *IEEE Trans. on Broadcasting*, Dec. 1991.
- [18] 정재길, 안치득, 이만섭, "미국의 전 디지털 HDTV", 주간기술동향 92-26, pp. 1-15, 한국전자통신연구소, 1992.
- [19] HDTV 특집, 전자공학회지 92-12, 1992.
- [20] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete Cosine Transform", *IEEE Trans. Computers*, vol. C-23, pp. 90-93, Jan. 1974.
- [21] J. Biemond, F. Bosveld and L. Lagendijk, "Hierarchical Sub-band Coding of HDTV in BISDN," ICASSP'90, pp. 2112-2116.
- [22] S. Mallat, "Multifrequency Channel Decompositions of Images and Wavelet models," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 2091-2110, Dec. 1989.
- [23] H. S. Malvar and D. H. Staelin, "The LOT : Transform Coding without Blocking Effects," *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. 37, no. 4, pp. 553-559, Apr. 1989.
- [24] 멀티미디어의 세계, 조사분석자료 93-01, 한국전자통신연구소, 1993.
- [25] 이병기, 강민호, 이종희, 광대역 통신 시스템, 교학사, 서울, 1992.
- [26] Future Technology Symposium Record, 18th Int'l TV Symposium, June 1993.
- [27] Yoshihiro Ishida, "Toward Integrated Digital Video for Next Generation", HDTV 최전선 - 21 세기의 방송/통신/컴퓨터/정보가전-, 멀티미디어 통신연구회, 조일신문사, 일본, 1993.
- [28] 박종훈, 이의택, "국내 광대역 영상서비스 전개방향 검토", '93 하계종합학술발표 논문집, 한국통신학회, 1993. 7.
- [29] Agreements on Profile/Level, ISO/IEC JTC1/SC29/WG11 N0489, July 1993. (S)

筆者紹介



安 致 得

1956年 8月 15日生

1980年 2月 서울대학교 공대 전자공학과(학사)

1982年 2月 서울대학교 대학원 전자공학과(석사)

1991年 7月 미국 University of Florida 전기공학과(박사)

1982年 12月 ~ 현재 한국전자통신연구소

1992年 7月 영상통신연구실장

주관심 분야 : 신호처리, 영상통신