

낮은 전송율의 전송로를 위한 영상신호의 감축방법에 관하여

金 聖 大

韓國科學技術院 電氣 및 電子工學科 助敎授

I. 서 론

TV신호의 디지털전송 가능성은 오래전부터 제시되어 왔으나 TV신호의 넓은 대역폭과 빠른 속도로 인하여 하드웨어 구현이 어려웠다. 그러나 요사이 반도체 기술의 혁신적인 발전으로 인한 값싸고 빠른 논리회로의 개발과 광섬유의 발달에 힘입은 광대역 전송방식이 가능해짐에 따라 아날로그전송보다 장거리전송에서 고도의 전송품질을 유지할 수 있고, 컴퓨터통신과 관련하여 영상정보의 컴퓨터처리가 유리한 디지털전송방식이 급격히 발전하고 있다. 특히 디지털전송방식은 CATV, 원거리영상회의의(Video-Teleconference), 영상전화(Picture-Phone) 및 방송중계등 여러가지 용도와 관련하여 매우 중요한 연구과제가 되고 있다.

우리가 통상 사용하고 있는 아날로그영상신호의 대역폭은 NTSC, PAL, SECAM 등 방식에 따라 5~8 MHz 정도이다. 이신호를 디지털방식으로 전송하기 위하여 8-bit PCM(Pulse-Coded-Modulation)으로 부호화하면 적어도 100Mbps(bits per second) 이상을 전송할 수 있는 전송로가 필요하게 된다. 이러한 막대한 량의 데이터를 한정된 전송속도를 갖는 전송로를 통하여 전송하기 위해서는 필수적으로 데이터량을 감축해야 한다. 현재까지 연구되어 있는 영상신호의 데이터 감축방법은 DPCM(Differential PCM)과 변환부호화(Transform Coding) 방법으로 대별된다. DPCM에 의한 감축방법은 현재 부호화 할 화소(Pel)의 실제값과 이미 부호화된 화소들로부터 얻을 수 있는 예측값과의 차이를 전송하는 방법으로 예측오차신호의 중복성(Redundancy)이 원래 신호의 중복성에 비해 적어지는 점을 이용하는 것이다. 변환을 이용한 감축방식은 원래의 영상을 적당한 변환방법을 사용하여 변환하게 되면 변환영역에서 일부 몇개의 변환계수만 큰 값을 가지고 대부분의 나머지 계수들은 작은 값을 가지게 된다는 점을 이용한다.

일반적으로 데이터감축방식의 선택은 전송로의 전송속도, 수신단에서 원하는 수신영상의 품질등에 달려있다. 방송중계등과 같이 고품질의 수신영상을 원하는 응용분야를 위해서는 45Mbps 정도의 속도를 가지는 전송로를 대상으로 한 DPCM에 의한 감축방식이 많이 연구·이용되고 있다. 12Mbps 정도의 낮은 속도의 전송로를 통하여 CATV 정도 품질의 영상을 송·수신하기 위한 방법도 많이 연구되고 있는데 이 경우에는 DPCM 방식에다 운동보상예측방법(Motion Compensated Prediction) 등을 이용하여 보다더 예측오차를 줄여서 데이터를 감축하는 방향으로 연구가 진행되고 있다. 원거리영상회의의 시스템, 영상전화등에 관한 연구에서 1.5Mbps 이하의 전송로를 대상으로 한 연구도 활발히 진행되고 있다. 이 분야에서는 화면간의 변화가 많지 않다는 특수성을 감안하여 여러가지 방법이 제안되고 있고 몇가지 방식은 실제로 상용화 되어 있는 실정이다. 현재 1.5Mbps 전송로를 이용할 수 있는 비디오코덱(Video-Codec) 시스템들 중에서 대표적인 제품으로는 CLI(Compression Lab. Inc)의 VTS1.5, NCE사의 NETEC-X1, 그리고 McM(GEC McMichael)의 VIDEO-CONFERENCE-CODEC 등이 있다. (표-1) 이들 제품에서 채택하고 있는 감축방식은 각기 달라서

표 1. 현재 상용화 되어 있는 1.5Mbps용 비디오코덱의 사양비교

	CLI(USA)	NEC(JAPAN)	McM(U. K.)
Video	1.3836Mbps	1.336 Mbps	1.28Mbps
Audio	111 Kbps	128 Kbps	64 Kbps
User data	20.4 "	8 "	64 "
T1 formatting	8 "	64 "	8 "
Control, sync, etc	21 "	8 "	128 "
Window size	368×480	455×525	320×286
Frame rate	15frames/s	30frames/s	N. A.

VTS1.5는 변환부호화방식을, NETEC-X1은 CRC (Conditional Replenishment Coding)방식을, McM 사의 제품은 CDC(Combinational Difference Coding)방식을 사용하고 있다. 이 시스템들 가운데서 CLI사의 DCT(Discrete Cosine Transform)을 이용한 VTS 1.5시스템의 성능이 가장 우수하다고 알려져 있다¹⁾

본 논고에서는 1.5Mbps정도의 낮은 전송속도를 가지는 전송로를 통하여 영상신호를 전송하기 위하여 사용되고 있는 감축방식에 대하여 살펴보기로 한다.

II. CDC(Combinational Difference Coding) 방식^{2,3)}

CDC는 공간상으로 인접한 화소들간의 상관관계와 시간축상으로 인접한 화소들간의 상관관계가 크다는 점을 이용하여 그림 1 과 같이 기본적으로 두개의 DPCM loop를 이용하여 영상데이터를 감축하는 방식이다.

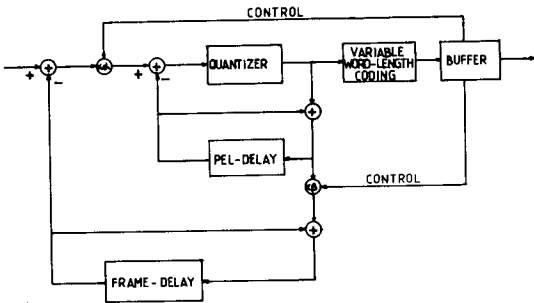


그림 1. CDC부호화기

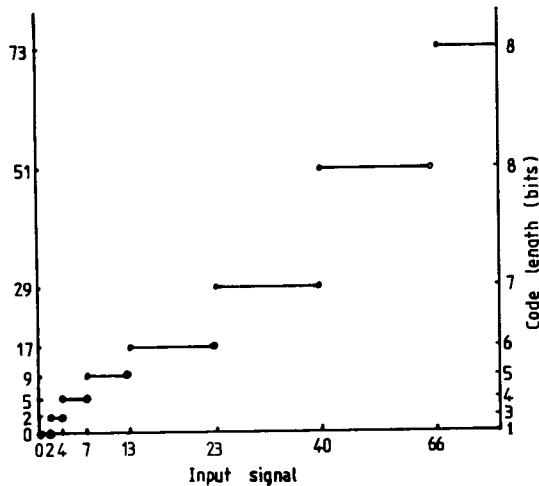


그림 2. CDC에서 사용되는 15-level 양자화기의 특성

Inter frame DPCM loop는 시간축상의 중복성

을 제거하는데 사용된다. 그림 1에서 보는 바와 같이 프레임메모리(frame memory)와 화소지연 소자(pel delay)를 이용하여 현재 부호화하고자 하는 화소를 예측한다. 이때 예측치와 실제값 사이에는 차이가 생기게 되는데 대부분의 경우 이 예측오차의 값은 상당히 작게 된다. 이 예측오차를 그림 2와 같은 특성을 가지는 양자화기(Quantizer)를 이용하여 일차적으로 데이터를 감축한다. 보다 더 데이터를 감축하기 위하여 양자화된 예측오차를 일종의 엔트로피 부호화방식인 VWC(Variable word-length coding) 부호화에 입력시킨다. 양자화된 예측오차 신호는 대부분의 경우 영일 확률이 크다. 그러므로 각 화소에 대하여 부호화를 하지 않고 작은 블록, 예를 들면 2x4 화소를 한 블록으로 하여 VWC를 적용시킨다. 이와같이 하여 부호화된 데이터는, 영상의 움직임 정도에 따라서 그 양이 계속 변하기 때문에 직접 전송로를 통하여 보낼 수가 없다. 그래서 buffer라는 완충장치를 두어 전송로를 통하여, 보내어지는데 데이터의 전송속도를 일정하게 해줄 필요가 있다. buffer의 상태는 영상의 움직임이 많은 화면이 계속될 경우, 전송속도보다 데이터발생량이 많을 경우, buffer의 용량을 초과할 우려가 있다. 이것을 overflow라고 하는데 이렇게 되면 데이터를 잃어버리게 된다. 그래서 적당한 방법을 사용하여 데이터의 발생량을 조절할 필요가 있다. CDC에서는 이를 위하여 buffer상태에 따라서 α 의 값을 조절한다. α 는 0과 1사이 에 있는 값을 취하는데 데이터량이 많아지면 α 의 값을 작게 하여 예측오차를 줄여서 양자화된 예측오차의 값이 영이될 확률을 높인다. 그와 반대일 경우 α 값을 크게 하여 보다 정확한 영상이 보내어지도록 한다. 그러나 이와같은 α 값의 조절에도 불구하고 계속 buffer의 양이 증가하면 field- 혹은 frame-subsampling을 행하여 데이터 발생량을 대폭 줄이는 방법을 사용한다. α 값이 작아져서 1/2정도보다 작아질 경우 blurring현상이 생기게 되는데 이것을 보상해 주기 위하여 β 값을 α 값의 변화에 따라 조절한다. β 값과 α 값의 관계는 그림 3과 같은 관계를 가지고 있다.

III. CRC (Conditional Replenishment Coding) 방식⁴⁾

CRC는 CDC에서처럼 매 화소마다 부호화를 하는 것이 아니라 이전 화면과 비교하여 변화가 있는 화소만 부호화하여 전송하는 방식이다. 그림 4에 CRC의 encoder부분을 나타내었다. 점선안의 부분은 inter-frame DPCM과 같은 형태이고 다만 이동검출기(MAD:

Moving Area Detector) 부분만이 추가로 있다는 것만이 다른 점이다.

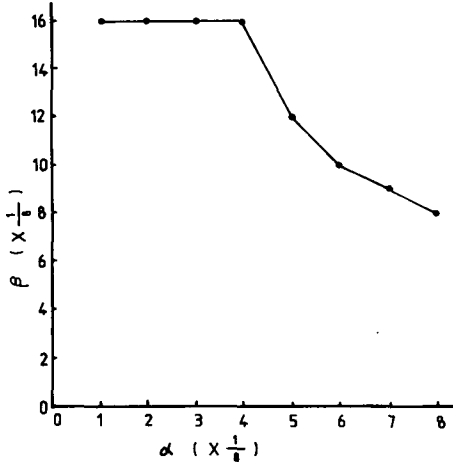


그림 3. CDC에서의 α 값과 β 값과의 관계

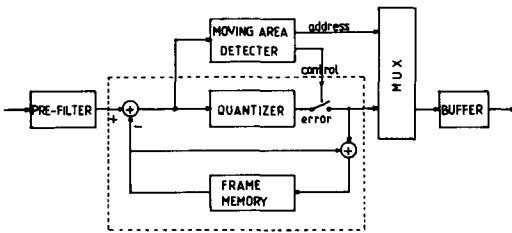


그림 4. CRC부호화기 송신단의 encoder

CRC 방식은 CDC에서와 마찬가지로 송·수신단에 프레임메모리가 있는데, 이동검출기는 이 프레임메모리에 들어있는 이전 화면과 입력되고 있는 현재의 화면과의 차이를 계산하여 움직임영역을 검출한다. 움직임영역 검출시 적당한 threshold 값을 이용하는데 그 차이가 이 threshold 보다 클 경우 해당 화소를 움직임영역에 포함시킨다. MAD에 의하여 검출된 움직임영역에 포함되는 화소들의 이전 화면에서의 값과 현재 화면에서의 값과의 차이를 양자화하여 그 화소의 위치 정보(address)와 함께 전송한다.

움직임영역이 작을 경우 앞의 방법으로 상당량의 영상데이터를 감축할 수 있지만 그렇지 못할 경우 위치 정보를 전송하는데 너무 많은 데이터량을 필요로 하게 된다. 이러한 점을 보완하기 위하여 움직임영역에 포함된 화소들 중에서 인접한 화소들끼리 묶는 clustering이란 방법을 사용한다. 그렇게 하면 매 화소마다 address를 보낼 필요가 없고 그 cluster의 address만 전송하면 된다. 경우에 따라 잡음의 영향으로 인하여 움직임영역내에 고립된 정지화소가 존재할 수도 있고, 정지영역내에 고립된 움직임 화소도 생길 수 있다. 이러한 고립된 화소들을 제거하면 위치전송에 필요한 데이터량을 줄일 수 있다. 이러한 방법을 bridging이라고 한다.

CRC의 경우도 데이터발생량이 불균일한데 발생된 데이터를 일정한 속도로 전송하기 위해서는 buffer가 필요하게 된다. CRC에서는 buffer 내에 있는 데이터량을 조절하기 위하여 MAD의 threshold 값을 조정하는 방법을 사용한다.

IV. DCT (Discrete Cosine Transform) 부호화방식^{15,16)}

변환부호화방식은 원래의 영상전체를 몇 개의 블록으로 분할한 후, 각 영역을 적당한 변환방법을 사용하여 변환한다. 일반적으로 변환을 하게 되면 대부분의 변환계수들은 영에 가까운 작은 값을 가지게 되고, 일부 몇 개의 변환계수들만이 큰 값을 가지게 되는데 이러한 것을 에너지집중현상(energy compaction)이라고 한다. 이때 적절한 방법으로 작은 값을 가지는 계수들을 버리고, 큰값을 가지는 계수들을 부호화함으로써 영상데이터를 감축할 수 있다.

변환방식중에서 가장 좋은 변환은 KLT(Karhunen-Loeve Transform)이다. 그러나 KLT는 영상의 covariance에 근거한 변환이기 때문에 TV영상과 같이 각 구역마다, 각 화면마다 covariance가 다른 경우매번 covariance를 구해야 한다는 난점이 있다. 뿐만 아니라 고속알고리즘이 없어 하드웨어나 소프트웨어의 구현이 용이하지 못하다.

영상처리에서의 응용을 위해 Haar 변환, Walsh-Hadamard 변환, Slant 변환, DST(Discrete Sine Transform), DCT(Discrete Cosine Transform) 및 DFT(Discrete Fourier Transform) 등 고속알고리즘을 갖는 준최적변환이 제안되었다. 그 중에서도 영상데이터가 stationary하고 지수형 상관계수를 갖는 일차 Mark-off process라고 가정할 경우 DCT는 에너지집중현상면에서 보았을 때 최적변환인 KLT와 가장 유사한 성능을 가지고 있고, 소프트웨어나 하드웨어의 구현이 가능하므로 변환을 이용한 데이터감축방법으로 많이 쓰이고 있다.

이차원함수 $f(j, k)$ 의 이차원 DCT는 다음과 같이 정의된다.

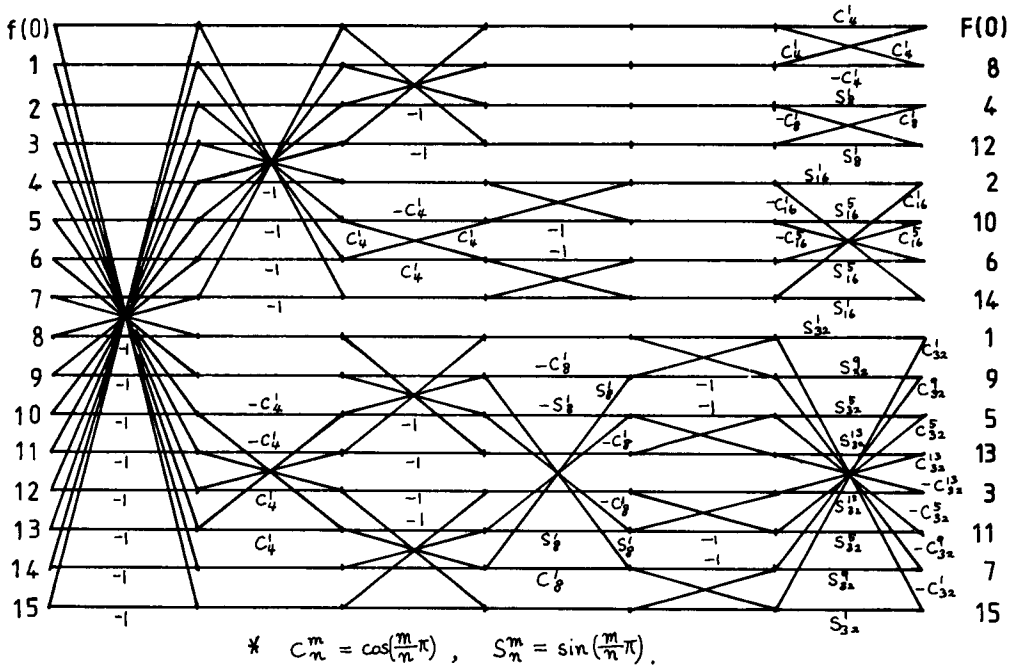


그림 5. Chen의 DCT변환 고속알고리즘 (N=16)

$$F(u, \nu) = \frac{4C(u)C(\nu)}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j, k) \cos\left[\frac{(2j+1)u\pi}{2N}\right] \cos\left[\frac{(2k+1)\nu\pi}{2N}\right]$$

윗 식에서 $j, k, u, \nu = 0, 1, 2, \dots, N-1$ 이고, $C(\omega)$ 의 값은 아래와 같다.

$$C(\omega) = \begin{cases} \frac{1}{\sqrt{2}} & (\omega = 0 \text{ 일때}) \\ 1 & (\omega = 1, 2, \dots, N-1 \text{ 일때}) \end{cases}$$

윗 DCT의 역변환은 다음과 같이 구해질 수 있다.

$$f(j, k) = \sum_{u=0}^{N-1} \sum_{\nu=0}^{N-1} C(u)C(\nu) F(u, \nu) \cos\left[\frac{(2j+1)u\pi}{2N}\right] \cos\left[\frac{(2k+1)\nu\pi}{2N}\right]$$

윗 식에서 알 수 있듯이 이차원 DCT변환은 두번의 일차원 DCT변환으로 구해질 수 있는데, 일차원DCT의 고속알고리즘은 Chen등에 의해서 제안되었다.

그림5는 Chen의 일차원 DCT의 고속알고리즘의 유통도를 그린 것이다. 역변환도 이와 비슷한 형태를 가지고 있다.

이러한 DCT변환을 이용하여 영상데이터를 감축하는 방법으로 가장 최근에 발표된 화상적용부호화기(SAC:

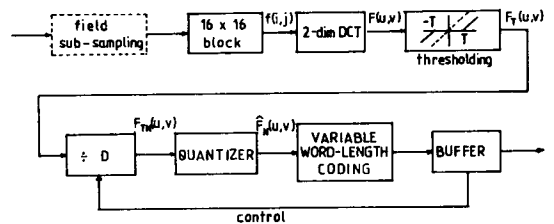


그림 6. DCT부호화기

Scene Adaptive Coder)를 소개한다. 이 SAC는 Chen에 의하여 발표되었는데 전체 시스템의 개략도는 그림6과 같다. 먼저입력영상을 작은 블록으로 쪼개어서 각 블록별로 DCT변환을 한 다음 적당한 threshold를 사용하여 큰 값을 갖는 계수들은 그대로 두고 작은 값을 가지는 계수들을 영으로 한다. 이 계수들을 그림7과 같은 순서로 부호화를 하는데 블록의 평균밝기를 나타내는 $F(0, 0)$ 는 9-bit로 PCM화하여 전송하고 그 이외의 계수들은 VWC(Variable Word-length Coding)를 하여 전송한다. VWC를 할때 영이 아닌 계수는 Huffman Coding을, 영인 계수는 Huffman의 run-length-coding을 한다.

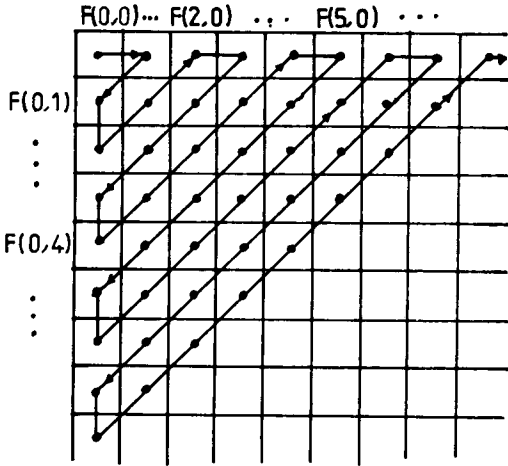


그림 7. DCT변환 계수를 부호화하는 순서

DCT부호화방식도 CDC나 CRC에서 처럼 데이터가 일정한 율로 발생하지 않는데 CDC나 CRC에서와 마찬가지로 buffer를 사용한다. buffer 조절은 그림6에서 보는 바와 같이 DCT변환계수를 thresholding 할 때 사용하는 threshold나 정규화계수인 D를 변화시킴으로 이루어질 수 있다.

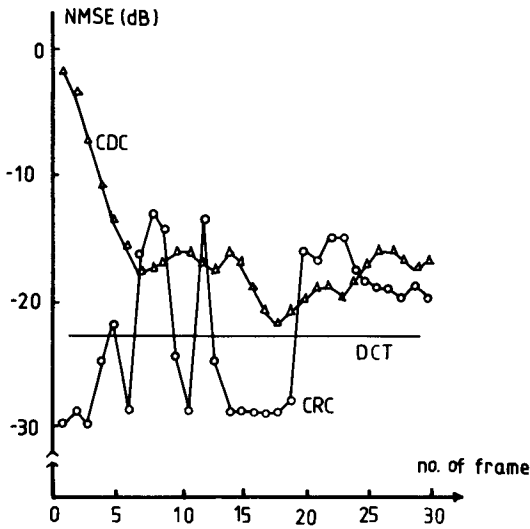


그림 8. CDC, CRC, DCT부호화기의 성능비교 (감축율=0.3bpp)

V. 세가지 방식에 대한 비교⁽⁷⁾

1.5Mbps 전송로를 통하여 영상을 전송하려면 적어도 평균한 화소당 0.3bit 정도로 감축하여야 한다. 그림8은 움직이는 연속적 영상인 CORNKITE에 대한 전술한 세가지 방식의 시간에 따른 NMSE(Normalized Mean Square Error) 변화를 나타낸 것이다. 사용한 감축율은 한 화소당 0.3bit가량이다. 여기서 NMSE는 다음과 같은 식으로 정의된다.

$$NMSE \triangleq 10 \log \frac{E\{(X - \hat{X})^2\}}{E\{X^2\}}$$

X : 원래영상 \hat{X} : 재생영상

그림8에서 알 수 있듯이 NMSE면에서 보면 CRC방식이 가장 우수하다고 할 수 있다. 그러나 영상의 변화가 심할 경우 NMSE값이 갑자기 커지는 것을 관찰할 수 있는데 이로 인하여 화면이 깜박거리는 현상이 일어난다. CDC는 다른 두 가지 방식에 비해서 성능이 떨어진다는 것을 알 수 있다. DCT를 이용한 부호화 방식의 NMSE가 거의 일정한 것은 DCT 변환이 영상의 통계특성의 변환에 민감하지 않음에서 기인한다.

하드웨어면에서 보면, 간단한 DPCM loop와 두개의 곱셈만 필요로 하는 CDC방식이 가장 간단하게 구현될 수 있고, DCT변화방식은 DCT를 실시간에 처리할 수 있도록 하는 하드웨어를 구현해야 하기 때문에 가장 복잡하다.

표 2는 위의 세가지 방식의 장단점을 비교한 것이다.

표 2. DCT, CRC, CDC부호화기의 장단점 비교

	Image quality	Buffer control	Hardware complexity
DCT codec	Good	Simple	High
CRC codec	Good	Complex	Medium
CDC codec	Not good	Very simple	Low

VI. 결 론

원거리영상회의 혹은 영상전화등에 이용가능한 영상신호의 데이터감축방법에 대하여 살펴보았다. 본 논고에서 언급된 세가지 방식은 현재 실제로 상품화되어 있는 1.5Mbps 전송로용의 Video Codec에서 실제로 사용되고 있는 것이다. 이들 방식들 사이에는 표 2에서 보는 바와 같이 서로 장·단점이 있지만 앞으로 반도체기술의 발전과 더불어 회로의 집적밀도가 높아지

게 되면 변환방법을 이용한 부호화 방식이 유리할 것으로 생각된다. 보다더 감축을 하기 위해서 변환방법과 DPCM방법을 결합하여 사용할 수도 있다.

參 考 文 獻

- [1] Dennis Bodson, et al, "The government's first step toward standardizing TV-motion gear," *Data Communication*, May, 1984.
- [2] Yasuda, et al, "Transmitting 4 MHz TV signals by combinational difference coding," *IEEE, COM-25*, pp. 508-516, 1977.
- [3] N. Mukawa, et al, "An interframe coding system for video teleconferencing signal transmission at a 1.5 Mbps rate," *IEEE, COM-32*, March, 1984.
- [4] B.G. Haskell, "Interframe coding of 525-line, monochrome television at 1.5 Mbit/sec," *IEEE, COM-25*, November, 1977.
- [5] W.H. Chen, et al, "A fast computational algorithm for the discrete cosine transform," *IEEE, COM-25*, September, 1977.
- [6] W.H. Chen, et al, "Scene adaptive coder" *IEEE, COM-32*, March, 1984.
- [7] 고종석, 김한수, 임재성, 장주욱, 김재균, "원거리 회의용 1.5Mbps 비디오코덱의 비교연구," 대한전자공학회 추계 학술대회논문집, 1984. *

1985 서울 國際 電子工學 심포지움 開催

본 학회에서는 다음과 같이 1985 서울 국제 전자공학 심포지움을 개최 하오니 이에 관심있으신 여러분의 많은 참여를 바랍니다.

행사명칭 : 1985 Seoul International Symposium on Electronics Engineering

목 적 : 한국, 일본, 중화민국의 세나라로 부터의 컴퓨터와 통신에 관련된 학술논문을 서울에서 발표함으로써 아세아에서의 이 분야 정보의 교환을 통해 전자산업의 발전을 촉진하고 나아가서 이 지역 국가의 번영에 기여하고자 한다.

주 최 : 대한전자공학회 (KIEE)

공동주최 : 일본전자통신학회 (IECEJ)

중국전기공학회 (CIEE)

후 원 : 체신부, 한국전기통신공사

협 찬 : IEEE 한국지부 대우전자(주)

삼성반도체통신(주) 국제전자(주)

금성반도체(주)

일 시 : 1985년 11월 15일(금)~16(토)

장 소 : 동방프라자 국제회의실