

# MPEG-4 동영상 부호화기를 이용한 조인트 소스/ 채널부호화에 관한 연구

\*이운문, \*신재은, \*손 원, \*\*이수인, \*\*이 현

\*경희대학교 전자정보학부 전파공학과, \*\*한국전자통신연구원 무선방송기술연구소  
\*king-moon@orgio.net

## Studies on Joint Source/Channel Coding Using MPEG-4 Video Codec

\*Woon Moon Lee, \*Jae Eun Shin, \*Won Sohn, \*\*Soo In Lee, \*\*Hyun Lee

\*Dept. of Radio Communication Engineering, School of Electronics and Information, Kyung Hee University  
\*\*Electronics and Telecommunications Research Institute

### 요 약

일반적인 통신 시스템의 소스 부호화기는 채널 잡음을 고려하지 않고 설계되며, 채널 부호화기는 소스 신호의 특성과 무관하게 채널 환경의 극복에만 중점을 두고 설계된다. 그러나 대역폭 제한적인 통신 환경에서 보다 효율적인 대역폭 사용을 위해서는 채널 환경에 따라 소스 부호율과 채널 부호율을 가변적으로 운용하여야 한다. 본 연구에서는 랜덤잡음 환경에서 동영상 신호를 전송함에 있어 채널 환경에 따라 한정된 채널 자원을 최적으로 할당하는 조인트 소스/채널 부호화 기법에 관하여 기술하였다. 소스 부호화로는 에러내성 기법을 적용한 MPEG-4 비디오 부호화기를 사용하였고, 채널 부호화는 RCPC 부호화기를 적용하였다. 본 논문에서는 여러가지 채널상태( $E_s/N_0$ )에 따른 율-왜곡특성에 대해 분석하여 소스부호율과 채널부호율 간에 최적의 상충점(Trade-off point)을 결정하였고, 주어진 채널 상태에 대한 목표부호율과 평균제곱오차(Mean Squared Error)의 관계에 대해서 기술하였다.

### I. 서 론

최근 들어 컴퓨터와 통신 기술의 비약적 발달로 인해 무선통신망 사용자들 사이에 대용량 멀티미디어 서비스에 대한 요구가 높아지고 있으나, 무선 통신의 주파수 자원은 한정되어 있는 상태이다. 따라서 한정된 주파수 자원을 얼마나 효율적으로 사용하는가 하는 것은 매우 중요한 문제이다. 기존의 통신 시스템의 경우 최악의 조건을 가진 채널 환경에 대해 목표로 하는 서비스 품질을 만족하도록 설계되었다. 그러나 이러한 시스템 설계는 최악의 채널 환경 이외의 대부분의 경우 대역폭 사용면에서 비경제적이라고 할 수 있다.

따라서, 이 논문에서는 제한된 대역폭에서 용량이 큰 멀티미디어 데이터를 전송할 때, 채널 환경에 따라 왜곡을 최소화하는 소스부호율과 채널부호율을 할당함으로써 시스템의 서비스 품질을 보장하고, 보다 효율적으로 대역폭을 사용할 수 있는 조인트 소스/채널 부호화 시스템에 대하여 논하였다.

샤논(Shannon)은 정보분리이론에서 무한대의 시퀀스 길이가 가정된 점 대 점 전송에서 소스부호화기와 채널부호화기는 각각 분리되어 최적화될 때, 소스 및 채널부호화기를 포함한 시스템이 최적화될 수 있다고 주장하였다. 그 동안

샤논이 주장해온 정보분리이론에 따라, 소스부호화기와 채널부호화기는 독립적으로 연구되었으나, 그의 이론이 가정된 '무한대의 시퀀스 길이' 조건과 '점 대 점 전송' 조건은 메모리 크기의 한계성 문제와 무선통신의 다중경로문제에 의하여 일반 통신시스템에서 만족될 수 없는 가정이라는 의견이 대두되기 시작하였다. 이에 따라, 1970년 후반부터 소스, 채널부호화기를 함께 고려하는 조인트 소스/채널 부호화 기술에 대한 연구가 시작되었다.

이 논문에서는 조인트 소스/채널 부호화의 최적화에 있어 다음과 같은 사항을 고려하였다. 채널은 랜덤잡음 환경을 적용했으며, 소스부호화기로는 MPEG-4 동영상 부호화기를 사용하였고, 채널부호화기로는 부호율 조절이 용이한 RCPC(Rate-Compatible Punctured Convolution) 부호화기를 사용하였다. 주파수 자원의 제한을 가정하기 위해 목표 비트율을 미리 정하고, RCPC 부호화기의 모부호율을 1/4로 하는 22개의 채널 부호율에 소스부호율을 조합시켜  $E_s/N_0$ 값에 따른 최적의 부호율을 결정하였다.

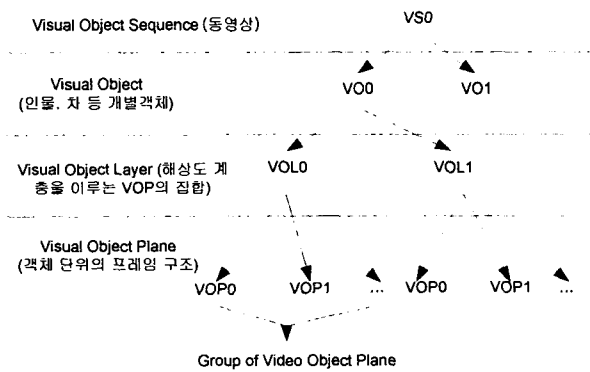
II장에서는 본 논문에서 적용된 MPEG-4 동영상 소스부호화기와 RCPC 채널부호화에 대해서 간략하게 소개했다. III장에서는 본 논문에서 적용한 조인트 소스/채널 부호화의 시스템 구조에 대해 논했으며, IV장에서는 모의 실험 결과, V장은 결론이 이어질 것이다.

## II. MPEG-4 동영상 부호화기와 RCPC 부호화기

이 장에서는 본 논문에 적용된 MPEG-4 동영상 소스부호화기와 RCPC 채널부호화기에 대해 논할 것이다.

### 1. MPEG-4 동영상 부호화기

MPEG-4는 디지털 저장 미디어와 디지털 동영상 통신을 위한 ISO/IEC의 동영상 압축 표준이다[4][5]. MPEG-4는 객체 기반 부호화와 계층적 부호화 등의 개념을 이용하여 부호화되며, 다음과 같은 선택스 구조를 갖는다.



<그림 1> MPEG-4의 선택스 계층

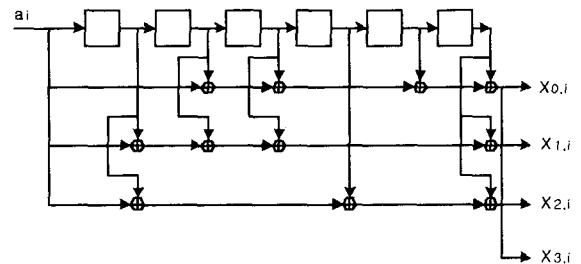
본 논문에서는 입력동영상 Stefan(QCIF)을 22가지의 서로 다른 초당 비트율(200 kbits/s ~ 582 kbits/s)로 부호화하여 비트열을 생성했다. 비트열은 채널부호화기와 랜덤잡음 채널을 거쳐 수신기에 전달되어지는데, 전송간에 발생되어지는 비트열의 에러는 복호 과정에 치명적일 수도 있다. 비트열은 단순 동영상의 시각적 데이터만을 가지고 있는 것이 아니라 특정한 구조를 유지하기 위해 헤더 성분과 시각적 데이터 성분으로 이루어져 있기 때문에 헤더성분에 에러가 발생하면 복호가 불가능 할 수도 있다.

송신기의 MPEG-4 부호화기에서 생성된 비트열은 <그림 1>과 같은 선택스 구조에 기반을 두게 되며, 수신기의 성공적인 복호화를 위해선 채널잡음이 더해진 비트열 구조도 같은 구조를 유지해야만 한다. 본 논문에서는 이러한 문제에 대해 어느 정도 대처하기 위해 MPEG-4 동영상 부호화기의 데이터 분할(Data Partitioning), 가역 가변장 부호화(Reversible Variable Length Coding)와 같은 에러 내성 기법을 사용했다.

### 2. RCPC 채널 부호화기

주어진 채널상태에서 왜곡을 최소로 하는 소스 부호율과 채널 부호율의 조합을 얻기 위해 소스부호화기의 출력동영상 비트열에 여러 가지 채널 부호율을 적용한다. 본 논문에서는 비교적 구현이 간단하고, 여러 가지 부호율을 얻는데 효율적인 RCPC 채널부호화기[2]를 사용하였다. 본 연구는 차후에 Eureka-147의 채널환경으로 확장을 시도할

예정으로, Eureka-147[3]에서 정의된 RCPC부호화기를 사용했으며, 24가지의 부호율 중에서 에러정정 능력이 떨어지는 높은 부호율(8/9, 8/10)들은 제외하고 22가지의 부호율에 대해 모의실험 하였다. RCPC의 모부호로는 구속장의 길이가 7이고, (133, 171, 145, 133)octal의 생성다항식을 갖는 컨볼루션 부호화기를 사용하였으며, 모부호로 사용한 컨볼루션 부호화기의 구조는 <그림 2>와 같다. 우리는 모부호에 의해 발생하는 코드워드를 모두 전송하지 않고, 평추어링을 사용하여 22가지의 채널부호율을 만들었다. 이때 평추어링 주기는 8로 하였으며, 평추어링 주기에 대한 모부호의 코드워드(32비트)에 대해 주어진 평추어링 패턴을 적용하여 22가지의 채널부호율( $R_c = 8/11 \sim 8/32$ )을 얻었다. 생성된 채널부호율을 소스부호화기의 출력 비트열에 적용하여 주어진 채널상태에서 왜곡을 최소로 하는 채널 부호율을 찾아냈다.



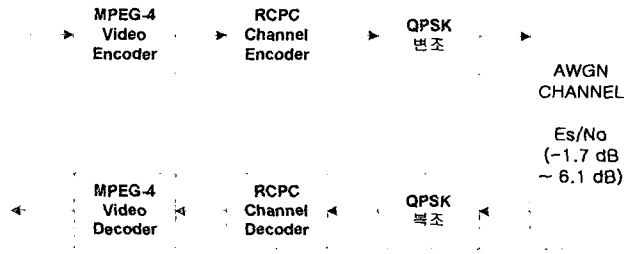
<그림 2> 컨볼루션 부호화기 (R=1/4)

채널복호화방법으로는 3bit-soft decision Viterbi 알고리즘을 사용하였다.

## III. MPEG-4 동영상 부호화기와 RCPC 부호화기를 적용한 조인트 소스/채널 부호화

지금까지의 대부분 무선통신서비스는 목표된 서비스 품질을 만족하기 위하여 어떤 채널환경에서도 목표 품질을 만족하도록 설계된 전송시스템을 운용하고 있다. 열악한 채널환경에서 주어진 품질을 만족하도록 설계된 전송시스템은 최악의 채널환경에서는 대역폭 사용효율 관점에서 적절한 전송방식을 사용하고 있지만, 대부분의 좋은 채널환경에서는 매우 비경제적인 전송방식을 사용하고 있다고 할 수 있다. 예를 들어, 주어진 C/N(반송파전력 대 잡음전력비)에서 열악한 채널환경에서는 목표 서비스 품질을 만족하기 위하여 최대 1 Mbps까지만 전송할 수 있으나, 거의 대부분의 경우에는 채널상태가 양호하여, 최대 10 Mbps까지 전송할 수 있다고 가정하자. 현재의 무선통신시스템은 항상 최대 1 Mbps까지 전송할 수 있도록 설계되기 때문에 채널상태가 양호한 경우에는 매우 비경제적으로 대역폭을 사용하고 있다고 할 수 있다.

본 논문에서는 동영상 전송관점에서 조인트 소스/채널 부호화를 적용했으며, <그림 3>과 같은 구조로 시스템을 구성하였다. 채널 상태에 따른 최적의 비트할당이 목적이기



<그림 3> 조인트 소스/채널 부호화의 시스템 구성도

때문에, 각 채널 상태(-1.7 dB ~ 6.1 dB)에 22가지의 부호율 조합을 적용하여 최소 왜곡을 갖는 부호율을 선택하였다. 울-왜곡 척도로 평균제곱오차(MSE, Mean Squared Error)값과 PSNR(Peak Signal to Noise Ratio)값을 사용하였으며, 각 왜곡 척도의 정의는 아래와 같다[6].

$$MSE(a, b) = \frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} [a(x, y) - b(x, y)]^2 \quad (1)$$

$a$ 는 입력동영상이고,  $b$ 는 출력동영상을 의미한다.  $N$ 과  $M$ 은 각각 동영상의 행 화소와 열 화소의 수를 가리키며,  $(x, y)$ 는 각 화소의 좌표에 대한 화소 값을 뜻한다.

$$PSNR(a, b)$$

$$= 10 \log_{10} \left[ \frac{255^2}{\frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} [a(x, y) - b(x, y)]^2} \right] \quad (2)$$

$$= 10 \log_{10} \left[ \frac{255^2}{MSE} \right]$$

#### IV. 모의 실험

본 모의실험에서는 소스부호화기로 MoMusys의 MPEG-4 코덱을 사용했고, MPEG-4에서 정한 움직임과 복잡도 정도가 큰 Class C급 Stefan(QCIF) 동영상을 소스로 채택하였다. 실험조건으로는 총 199 프레임, 프레임율 30 Hz, 하나의 I 프레임에 198개의 P 프레임을 갖는 형태, 즉 IPPP 형태로 부호화했다. 또한, 비트열 손실을 최소화 하기 위해 데이터 분할 기법과 가역 가변장 부호화를 사용하였다. 채널부호화로는 <그림 2>의 구조를 갖는 컨볼루션 부호화기에 펄쥘어드를 사용하여 다양한 코드율을 얻는 RCPC를 사용하였다. 모의실험에서 사용한 부호율은  $R_c = 8/(8+PI)$ 이며,  $PI$ (puncturing index) = 3~24이다. 복호화로는 3bit-soft decision Viterbi 알고리즘을 사용하였다. 또한 QPSK 변조방식을 사용하여 랜덤잡음 채널에서 모의 실험하였다.

소스부호화기와 채널부호화기에 할당되는 목표 비트율은 300 kbits/s 이고, 채널부호율에 따른 22개의 부호율 조합에 대해 Es/No값을 -1.7dB부터 0.3dB 주기로 6.1dB의 범위까지 모의 실험을 수행하였다.

#### 1. Es/No 값에 따른 울-왜곡 특성

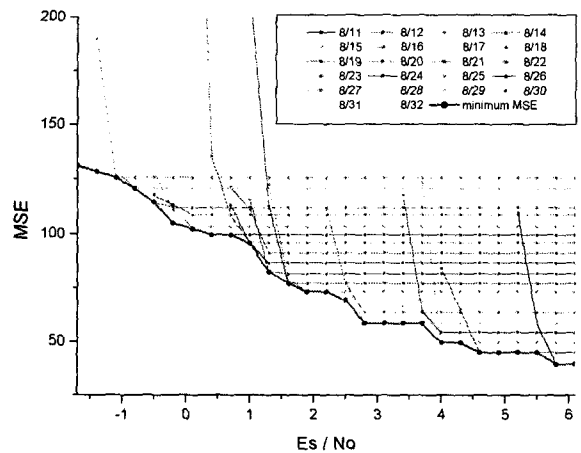
목표비트율(800 kbits/s)이 정해지면, 식 (3)을 이용하여 <표1>과 같은 비트율 조합을 정의할 수 있게 된다. 여기서  $R_T$ 는 목표비트율 이고,  $R_S$ 와  $R_C$ 는 각각 소스비트율과 채널부호율을 나타낸다.

$$R_T = R_S + R_C \quad (3)$$

<표 1> 목표비트율에 대한 부호율 조합

No	채널부호율	소스비트율	No	채널부호율	소스비트율
1	8/11	582 kbits/s	12	8/22	291 kbits/s
2	8/12	533 kbits/s	13	8/23	278 kbits/s
3	8/13	492 kbits/s	14	8/24	267 kbits/s
4	8/14	457 kbits/s	15	8/25	256 kbits/s
5	8/15	427 kbits/s	16	8/26	246 kbits/s
6	8/16	400 kbits/s	17	8/27	237 kbits/s
7	8/17	375 kbits/s	18	8/28	229 kbits/s
8	8/18	356 kbits/s	19	8/29	221 kbits/s
9	8/19	338 kbits/s	20	8/30	213 kbits/s
10	8/20	320 kbits/s	21	8/31	206 kbits/s
11	8/21	305 kbits/s	22	8/32	200 kbits/s

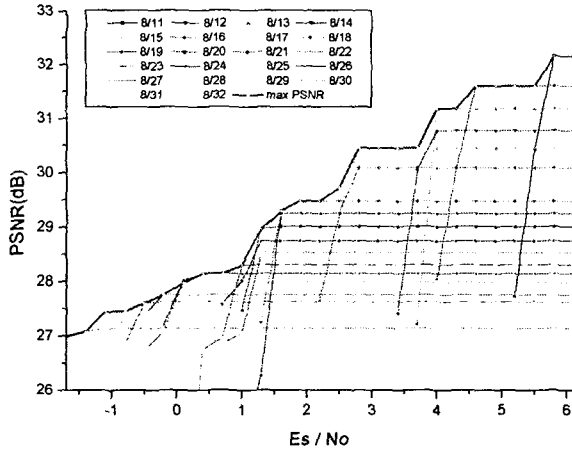
<그림 4>는 Es/No 값에 따른 평균제곱오차값을 나타낸다. 여러 가지 채널 상태에 대해 각 부호율 조합이 어떠한 왜곡 특성을 갖느냐를 나타내는 것으로서, 각 채널 상태마다 최소의 왜곡을 갖는 부호율의 조합을 발견할 수 있다.



<그림 4> Es/No 값에 따른 평균제곱오차

최소평균제곱오차(minimum MSE)로 나타내어지는 선은 최소 왜곡을 갖는 부호율 들의 집합들로서, 주어진 채널상태(Es/No)에 대해 최소의 왜곡을 갖는 최적의 비트할당 조합을 나타낸다.

<그림 5>는  $E_s/N_0$ 값에 따른 복호된 동영상의 평균 PSNR을 나타낸 것으로, 그림에서 보여지는 최대 평균 PSNR(max PSNR)은 <그림 4>의 최소평균제공오차에서와 동일한  $E_s/N_0$ 에 대해 동일한 부호율 조합을 갖게 된다. <그림 4>와 <그림 5>에서 x축( $E_s/N_0$ 의 축)에 평행한 직선을 그리는 구간들은 각각 같은 평균제공오차와 PSNR을 갖는 구간들이다. 이 구간들은 소스부호화로 인한 왜곡만이 존재하는 구간으로서 더 이상의 채널잡음에 의한 어려움이 없는 부분을 나타낸다. 즉, 주어진 채널부호화율로 100% 비트보호가 가능한 구간을 의미한다. 이러한 경우에는 소스부호율에 좀 더 크게 비중을 두어 보내고자 하는 동영상의 PSNR 값을 높이는 과정이 요구된다. 결과적인 소스부호율과 채널부호율의 최적의 조합은 <표 2>에서 볼 수 있다.



<그림 5>  $E_s/N_0$  값에 따른 PSNR

<표 2> 각  $E_s/N_0$ 에 대한 최적의 부호율 조합

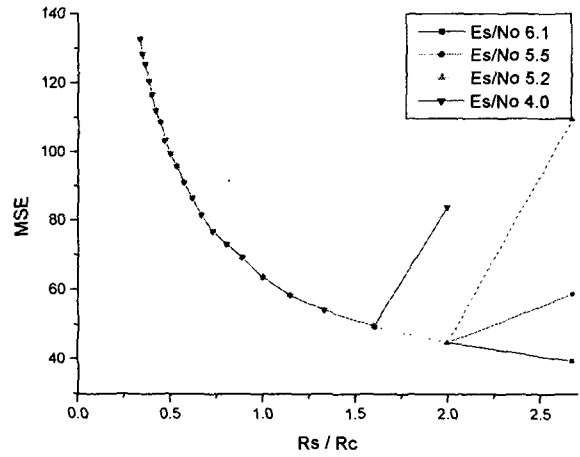
$E_s/N_0$	채널부호	소스비트율	2.2dB	8/17	375 kbits/s
-1.7dB	8/31	206 kbits/s	2.5dB	8/16	400 kbits/s
-1.4dB	8/31	206 kbits/s	2.8dB	8/15	427 kbits/s
-1.1dB	8/28	229 kbits/s	3.1dB	8/15	427 kbits/s
-0.8dB	8/27	237 kbits/s	3.4dB	8/15	427 kbits/s
-0.5dB	8/27	237 kbits/s	3.7dB	8/15	427 kbits/s
-0.2dB	8/25	256 kbits/s	4.0dB	8/13	492 kbits/s
0.1dB	8/24	267 kbits/s	4.3dB	8/12	533 kbits/s
0.4dB	8/23	278 kbits/s	4.6dB	8/12	533 kbits/s
0.7dB	8/22	291 kbits/s	4.9dB	8/12	533 kbits/s
1.0dB	8/21	305 kbits/s	5.2dB	8/12	533 kbits/s
1.3dB	8/19	338 kbits/s	5.5dB	8/12	533 kbits/s
1.6dB	8/18	356 kbits/s	5.8dB	8/11	582 kbits/s
1.9dB	8/17	375 kbits/s	6.1dB	8/11	582 kbits/s

최적의 부호율의 한 예로, <표 2>에서의 5번 부호율 조합 (427 kbits/s의 소스비트율과 8/15의 채널부호율)은 채널상

태( $E_s/N_0$ )가 2.8dB ~ 3.7dB 구간에서 최소왜곡을 갖는 것을 볼 수 있다. 이 구간의 채널상태에서는 다른 어떤 부호율의 조합보다 우수한 성능을 나타내게 되는 것이다.

## 2. 부호율에 따른 율-왜곡 특성

<그림 6>은  $R_s/R_C$ 와 평균제공오차의 관계를 나타낸 그림이다. x축은 부호비율(소스비트율/채널부호율)  $R_s/R_C$ 를 나타내고, y축은 평균제공오차를 나타낸다. x축의 값 1.0을 기점으로 왼편은 <표 1>에서 7~22 조합의 값을 나타내고, 오른편은 1~5 조합의 값을 나타낸다.



<그림 6> 여러 가지 채널 환경에 대한 평균제공오차 vs  $R_s/R_C$

<그림 6>을 보면, 4가지의 채널상태에 대하여 소스부호율 비율이 증가할수록 왜곡의 크기가 점차 줄어들다가 각 채널의 특성에 따라 특정  $R_s/R_C$ 에서부터 다시 왜곡이 증가하는 것을 볼 수 있다. 이처럼 율-왜곡 곡선이 하향곡선을 그리다 상승곡선으로 변경되는 점이 해당 채널상태에 대한 최소왜곡이 되는 것이다. 예를 들어,  $E_s/N_0 = 4.0$  dB의 경우  $R_s/R_C = 1.6$  부근에서 최소의 왜곡을 알 수 있다.

## 3. $E_s/N_0 = 1$ dB에서의 율-왜곡 특성 비교

<그림 7>과 <그림 8>은  $E_s/N_0 = 1$ dB에서 서로 다른 부호율을 가지고 부호화된 동영상이다.

<그림 7>은  $E_s/N_0 = 1$ dB에서 최소의 왜곡을 갖는 동영상으로 <표 1>에서 11번 조합에 의해 부호화되었다. <그림 8>은 같은 채널상태이긴 하지만 다른 부호율인 10번 조합으로 부호화된 동영상이다.

<그림 8>의 동영상은 <그림 7>의 동영상보다 소스부호화에 비트할당을 많이 하지만, 특정 채널 상태에 대하여 적절한 채널 부호율을 할당하지 못하여 더 열화된 동영상을 얻는 것을 볼 수 있다. 여기서, <그림 7>이 최적의 비

트할당을 이루는 조인트 소스/채널 부호화에 의한 동영상에 해당하는 것이다.

우수한 성능을 가지는 것으로 분석되었다.

### 참고 문헌

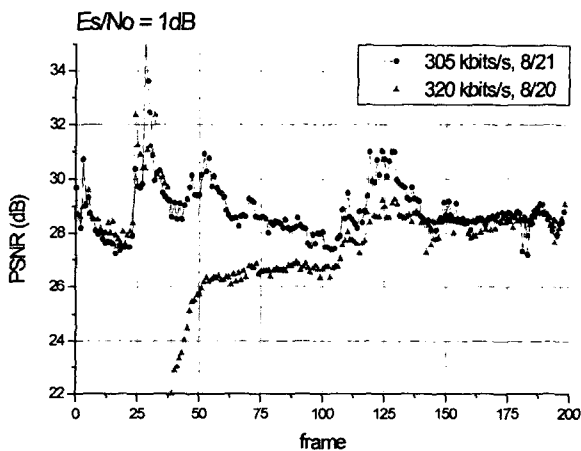


<그림 7> Es/No = 1dB  
(305 bits/s, 8/21)  
PSNR = 28.75 dB



<그림 8> Es/No = 1dB  
(320 bits/s, 8/20)  
PSNR = 27.48 dB

<그림 9>는 두 동영상의 PSNR을 비교이다. 10번 조합의 동영상은 35프레임 이후에 큰 비트 손실에 의하여, 갑작스런 화질 열화(PSNR 손실)가 일어남이 보여진다. 반면에 조인트 소스/채널 부호화에 의한 11번 조합의 동영상은 비교적 균형적인 화질을 유지함을 볼 수 있다.



<그림 9> Es/No = 1dB에서의 프레임 vs PSNR

### V. 결론

본 논문에서는 MPEG-4 동영상 전송을 목적으로 대역폭 제한적인 통신 채널에서 대용량의 멀티미디어 정보를 전송함에 있어 소스부호화기와 채널부호화기의 부호율을 채널환경에 따라 한정된 채널 자원을 최적으로 할당함으로써 서비스 품질과 대역폭 사용 효율 면에서 최적의 성능을 갖도록 하는 조인트 소스/채널 부호화 시스템에 대해서 논하였다. 여러 가지의 소스부호율과 채널부호율의 조합을 다양한 채널상태에 적용시킴으로써 특성화된 각 채널상태마다 조인트 소스/채널 부호화의 최적 부호율을 할당할 수 있었으며, 조인트 소스/채널 부호화 기법을 적용한 경우의 부호율의 조합이 그렇지 않은 경우에 비하여

[1] Gene Cheng, Avideh Zakhor, "Bit Allocation for Joint Source/Channel Coding of Scalable Video", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 9, No 3, March 2000.

[2] Joachim Hagenauer, "Rate-Compatible Punctured Convolutional Codes(RCPC Codes) and their Applications", IEEE TRANSACTIONS ON COMMUNICATIONS, VOL 36, No 4, April 1988.

[3] "Radio broadcasting systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers" ETS 300 401, May 1997

[4] Fernando Pereira, Touradj Ebrahimi, "The MPEG-4 BOOK", IMSC, 2002.

[5] 三木 彌一 著, 고성재 · 김종욱 역, "다양한 영상 · 음성을 자유자재로 부호화하는 MPEG-4의 세계", 대영사, 1999.

[6] King N. Ngan, Chi W. Yap, Keng T. Tan, "Video Coding for Wireless Communication Systems", MARCEL DEKKER, INC. 2001.