

# H.263+을 위한 MAP기반의 Joint Source-Channel Coder 설계

송 호 현, 최 윤 식

연세대학교 전기. 컴퓨터 공학과

전화: 02-361-2774 / 핸드폰: 017-303-4501

## Design of Joint Source-Channel Coder for H.263+ by MAP estimation

Ho-Hyon Song, Yoon-sik Choe

Dept. of electrical and computer, Yonsei University

E-mail: hodai@image3.yonsei.ac.kr

### Abstract

In this paper, We try to design combined source-channel codec that is compatible with video coding standards. This MAP decoder is proposed by adding semantic structure and semantic constraint of video coding standards to the method using redundancy of the MAP decoders proposed previously. Then, We get the better performance than usual channel coder's.

### I. 서론

요즘 통신에 대한 수요가 증가함에 따라 여러환경 하에서의 정확한 정보의 전송에 대한 연구가 이루어지고 있다. 특히 단순한 음성의 전송에만 치중되었던 통신이 일반인의 멀티미디어의 전송에 대한 관심의 증가로 인해 통신상에서의 영상의 전송도 중요시되고 있다.

하지만, 음성전송의 경우와는 달리 영상의 전송에 있어서 채널의 오류는 치명적인 영향을 미칠 수가 있다. 만약 고정부호가 사용된다면, 채널의 오류에 의한 영향은 단지 하나의 코드워드에만 영향을 주지만, 요즘의 영상 표준은 가변부호를 기본으로 하기 때문에 단 한 비트의 채널오류에 의한 영향이 다음 재동기화가 일어나기 전까지 전파되어 안정적인 영상을 얻기가 힘들어진다.

이런 이유로 채널오류에 강인하고 이를 정정할 수 있는 채널 복호화기의 중요성이 높아지고 있다. 하지만 채널오류를 고려한 채널복호기는 단순히 채널상의 오

류확률에 의한 오류를 정정할 수 있을 뿐 소스내의 각 비트 사이에 존재하는 상관도(correlation)를 고려한 오류의 정정은 고려되지 않는다. 즉, 소스 부호화기는 소스의 중복성을 최대한 줄여서 전송하고, 채널 부호화기는 채널오류에 의한 영향을 줄이기 위해 부가정보를 추가함으로써 이루어진다. 이는 Shannon에 의한 separation theorem에 의한 것이나, 요즘의 실제 시스템의 경우 이러한 분리는 가능하지 않다. 또한, 소스가 확률적이라는 가정 하에서 시작하는 소스부호화도 소스의 모든 중복 정보를 없애지는 못한다.[7]

이런 이유로 채널과 소스의 관계를 분리시키지 않고 서로 연관시켜서 더 나은 성능을 얻으려는 연구들이 이루어지고 있는데, 채널오류로 인해 영향을 받은 데이터를 정정할 때, 올바르게 복호된 소스의 부가정보를 이용하려는 연구들이 이에 속한다.[1]-[5]

본 논문에서는 채널의 오류확률과 소스의 전이확률 분포를 이용하여 채널 오류에 의해 손상된 정보를 올바르게 정정하는 방법을 제시하고자 한다. 기존의 논문들에 의하면 채널 에러에 의해 동기를 잃는 것을 막기 위해 고정부호의 경우로 제한하여 소스 전이확률을 이용하거나,[1][3] 가변부호의 경우도 픽셀값을 코딩하기 위한 DPCM 코드나 움직임 벡터를 허프만 코드로 나타내는 간단한 경우의 부가정보만을 채널 오류 정정을 위해 사용하였다.[2][4][5]

하지만, 여기서는 가변부호인 H.263+이라는 영상압축 표준에 적용가능한 복호기를 설계하였다.

## II. 영상 시스템에서의 redundancy

최근 들어 채널과 소스의 관계를 분리시키지 않고 서로 연관시켜 더 정확한 정보를 얻고자 하는 joint source-channel coding에 대한 연구에 관심이 집중되고 있다. 이는 채널에 오류가 존재하고 채널의 대역폭이 제한된다는 상황에서 무선 통신의 중요성이 높아지고 있기 때문이다. 과거의 경우 소스 부호화와 채널 부호화는 분리되어 설계되었다. 이는 Shannon에 의한 separation theorem[6]에 의한 것으로 최근 들어 Shannon의 이론은 모든 채널에서 성립하지는 않는다는 것이 증명되었다.[7] 특히 제한 조건이 많은 이동통신 채널의 경우 성립하지 않는 것으로 알려져 있다.

[1]에 의하면 DPCM 시스템에 대해서 양자화에 의한 소스의 output에는 redundancy가 존재한다. 마찬가지로 다른 부호화 시스템의 경우 아주 조금의 부호화가 존재하더라도 소스 부호화의 output에는 redundancy가 존재하게 된다.[3]

## III. 동영상에서 에러의 영향

일반적으로 많은 low-bit rate video coding (MPEG-1, 2, H.263등)들은 시간적 중복성과 공간적 중복성을 줄이는 방법을 통하여 압축을 하게 된다. 이렇게 해서 줄여진 데이터들에 대해서 허프만 코드와 같은 가변부호 방법을 사용하여 더 압축된 데이터를 만들어 낸다.

이처럼 동영상압축과정에서 흔히 쓰이는 DCT나 움직임 보상 알고리즘에는 가변부호를 사용한다. 따라서 일단 이러한 가변부호에 발생한 오류는 복호화과정에서 이후의 데이터에까지도 영향을 미쳐 복호기의 재동기가 이루어지기 전까지 다른 블록들에 영향을 미치고 이렇게 공간적으로 전파된 오류는 움직임 보상에 의해 다음 픽처로 전파된다.

H.263에서는 16×16크기의 매크로 블록들을 그림의 한 열 단위로 모아 GOB라고 하고 MPEG등에서는 비슷한 개념으로 슬라이스라는 단위를 두고 있다. GOB, 혹은 슬라이스 계층내의 데이터 열은 대부분 가변부호로 이루어져 있고 복호기의 재동기는 픽처 개시부호(PSC)나 GOB 개시부호(GBSC) 혹은 슬라이스 개시부호(SSC)에 의하여 이루어진다. 따라서 GOB나 슬라이스헤더보다 하위계층에서 에러가 발생하면 그 영향이 GOB내 또는 슬라이스내에서 전파가 이루어 질 수 있다. 만약 GOB나 슬라이스를 사용하지 않는다면 에러에 의한 영향은 에러가 발생한 이후부터 그림의 끝까지 전파가 된다.

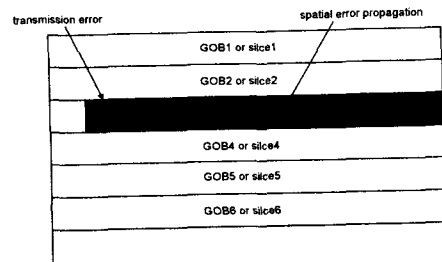


그림 1. 픽처내에서 공간적 에러의 전파

그림 1은 영상의 한 매크로블록에서 가변부호의 혼란으로 인해 오류가 전파되는 공간적 전파과정을 보여준다.

영상 압축 표준들은 앞에서 얘기한 DCT같은 공간적 중복성 제거뿐만이 아니라 움직임 보상 같은 시간적 중복성도 이용을 한다. 그러므로 한 프레임의 그림에서 오류가 발생할 경우 그 그림 내에서 오류가 전파될 뿐 아니라 다음 그림에도 영향을 주기 때문에 단 하나의 오류도 영상의 화질에 큰 영향을 미치게 된다.

## IV. MAP decoding

이러한 이유로 복호기에서 데이터를 복호할 때, 단순히 채널 오류에 의한 채널 복호화만이 아니라 소스내에 존재하는 redundancy를 이용함으로써 더 정확한 데이터의 복원을 기대할 수 있음을 알 수 있다.

크기  $N$ 의 alphabet  $J_N = \{0, 1, \dots, N-1\}$ 을 가지는 stationary stochastic process  $\{I_n\}_{n=1}^{\infty}$ 이 모든  $n > 1$ 에 대해 다음 조건을 만족하면, discrete Markov source라고 한다.

$$\Pr\{I_n = i_n | I_1^{n-1} = \mathbf{i}_1^{n-1}\} = \Pr\{I_n = i_n | I_{n-1} = i_{n-1}\} \quad (1)$$

$$i_k \in J_N, k = 1, 2, \dots, n,$$

$$\begin{aligned} I_1^{n-1} &= (I_1, I_2, \dots, I_{n-1}), \\ \mathbf{i}_1^{n-1} &= (i_1, i_2, \dots, i_{n-1}) \end{aligned} \quad (2)$$

$\Pr\{I_n = i_n | I_{n-1} = i_{n-1}\}$ 은 stationarity의 조건으로부터  $n$ 에 독립적이며 Markov source는 소스 전이확률값으로 표현된다.

$$P(i_n | i_{n-1}) = \Pr\{I_n = i_n | I_{n-1} = i_{n-1}, i_n, i_{n-1} \in J_N\} \quad (3)$$

간단한 discrete Markov source의 예는  $N=2$ 인 binary symmetric Markov source로서

$$P(i_n | i_{n-1}) = \begin{cases} p & \text{if } i_n = i_{n-1} \\ 1-p & \text{if } i_n \neq i_{n-1} \end{cases}, \quad (4)$$

$$i_n, i_{n-1} \in J_2$$

이다.

이러한 Markov source  $\{I_n\}_{n=1}^{\infty}$ 는 source와 같은 alphabet  $J_N$ 을 가지는 discrete memoryless channel을 통과한다. 채널의 output을  $\{J_n\}_{n=1}^{\infty}$ 이라 하면, 채널의 확률은 채널 오류확률값으로 표현된다.

$$Q(j_n|i_n) = \Pr\{J_n=j_n|I_n=i_n\}, \quad (5)$$

$i_n, j_n \in J_N$

간단한 discrete memoryless channel의 예는  $N=2$ 인 binary symmetric channel로서.

$$Q(j_n|i_n) = \begin{cases} 1-\varepsilon & \text{if } j_n=i_n \\ \varepsilon & \text{if } j_n \neq i_n \end{cases}, \quad (6)$$

$i_n, j_n \in J_2$

이다.

이러한 가정 하에 discrete memoryless channel을 통해 전송되는 discrete Markov source의 detection을 고려할 때 오류의 확률을 최소로 만드는 maximum a posteriori(MAP) criterion을 생각하고자 한다. 이전의 연구들도 마찬가지로 오류의 확률을 최소화하는 MAP 복호기를 가정하였다. 이때 채널 오류로부터의 보호를 위해 Markov source의 연속적인 심볼사이의 redundancy를 이용하였다. 여기서 다른 MAP detection은 sequence MAP detection으로 복호기에서 시퀀스를 받았을 때, 전송되었을 가장 확률이 큰 시퀀스를 결정하는 것이다.

소스의 크기  $N$ , 소스 전이확률값  $P$ , 초기 소스확률값  $P_1$ , 채널 에러확률값  $Q$ 와 받아진 시퀀스  $J_1^n = j_1^n$ 이 주어졌을 때 sequence MAP detection은 다음 식을 만족한다.

$$\hat{i}_1^n = \arg \max_{i_1^n \in J_N^n} \Pr\{I_1^n = i_1^n | J_1^n = j_1^n\} \quad (7)$$

(7)식은 Bayes' Theorem에 의해 다음과 같이 쓸 수 있다.

$$\hat{i}_1^n = \arg \max_{i_1^n \in J_N^n} \frac{\Pr\{J_1^n = j_1^n | I_1^n = i_1^n\}}{\Pr\{J_1^n = j_1^n\}} \cdot \Pr\{I_1^n = i_1^n\} \quad (8)$$

$J$  항은  $\hat{i}_1^n$ 을 최대화하는데 영향을 주지 못한다고 가정한다.

$$\hat{i}_1^n = \arg \max_{i_1^n \in J_N^n} \Pr\{J_1^n = j_1^n | I_1^n = i_1^n\} \cdot \Pr\{I_1^n = i_1^n\} \quad (9)$$

채널이 memoryless이고, conditional probability와 source의 Markovian property를 적용하면, 다음과 같

이 표현할 수 있다.[3]

$$\hat{i}_1^n = \arg \max_{i_1^n \in J_N^n} \left[ \sum_{k=2}^n \log[Q(j_k|i_k)P(i_k|i_{k-1})] + \log[Q(j_1|i_1)P_1(i_1)] \right] \quad (10)$$

위의 형태에서 sequence MAP detection은 Viterbi algorithm과 비슷한 형태를 보임을 알 수 있다.

## V. 제한하는 MAP decoding

기존의 논문에서 제한한 방법들은 training 시퀀스를 이용하여 한 심볼에서 다른 심볼로 전이하는 과정을 확률로 정의하였다. 하지만 이러한 가정은 모든 영상에 대해 적용 가능하다고는 볼 수 없다. 이런 이유로 여기서는 모든 영상에도 보편적으로 적용할 수 있는 소스 전이확률을 제안하고자 한다.

H.263+[8]과 같은 영상 압축 표준들은 헤더 정보와 데이터 정보로 나누어진다. 보통 중요정보인 헤더 정보는 각 시퀀스의 처음에 존재하고 고정부호이다. 하지만 데이터 정보는 가변부호로 이는 허프만 코드로 이루어져 있다. 허프만 코드는 실험적으로 많이 나오는 심볼에 짧은 비트를 할당하고 적게 나오는 심볼에 많은 비트를 할당하여 전체적으로 적은 비트를 사용하도록 하는 방법이다.

예를 들어 그림 2와 같이 1, 01, 00이라는 허프만 코드를 사용하는 코드가 존재한다고 가정한다.

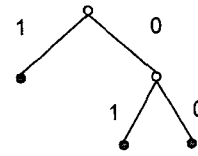


그림 2. 허프만 코드의 예

허프만 코드의 정의에 의해 1은 1/2의 확률을 01과 00은 각각 1/4의 확률을 가진다. H.263에는 P-픽처의 경우 Header, COD, MCBPC, CBPY, DQUANT, MV, INTRADC, TCOEFF, ESCAPE라는 허프만 코드가 존재하는데 각각의 경우 다른 확률분포를 가진다. 이러한 허프만 코드의 확률분포를 소스 전이확률로 이용하였다. 즉, 심볼의 길이에 반비례하여 확률을 할당하였다. 이와 함께 H.263내의 문법적인 구조와 문법적인 제한조건을 이용하여 더 정확한 복호화를 시도하였다. 문법적인 구조는 H.263의 각각의 심볼들이 올 수 있는 순서를 제한하여 더 정확한 복호화를 유도하는 방법이고, 문법적인 제한조건은 다음의 다섯가지 제한을 사용한다.

(1)COD가 '0'이면, 다음은 MCBPC가 온다.

- COD가 '1'이면, 다음은 COD가 온다.
- (2)COD의 개수는 한 GOB당 11개이다.
- (3)MV는 연속적으로 두 개만이 올 수 있다.
- (4)INTRADC는 연속적으로 6개를 넘을 수가 없다.
- (5)한 블록당 TCOEFF의 개수는 64개를 넘을 수 없다.

### VI. 실험 결과 및 고찰

실험은 H.263+의 base line을 기준으로 하였으며, 기존의 Viterbi algorithm과 같은 조건하에서 실험을 하였다. 아래 그림의 경우처럼 H.263에 의한 output에 1/3 convolutional coder를 적용하였고 채널 오류를 삽입한 후 채널 복호기를 Viterbi algorithm의 경우와 제안하는 MAP의 경우로 나누어 실험하였다.



그림 3. Block diagram

소스는 GOB layer option을 추가하여 오류의 전파를 한 GOB로 제한하였고, burst error하에서 성능 비교를 하였다.

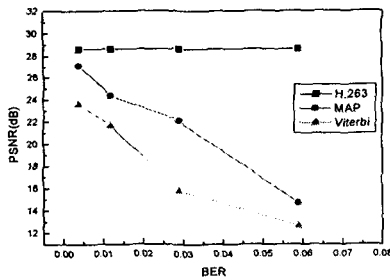


그림 4. Foreman 결과 영상의 PSNR

위의 결과에서처럼 burst error가 존재하는 경우 Viterbi 복호기에 의한 영상은 움직임 벡터와 DCT정보가 정확하게 복원되지 않아 성능이 떨어짐을 알 수 있지만, 제안하는 MAP의 경우는 더 정확하게 복원함을 알 수 있다.

### VII. 결론

Viterbi 복호기는 한 비트가 다른 비트로 변화할 때의 가능성을 이용해 오류로부터의 영향을 줄이는 복호기이다. 하지만 Viterbi 복호기는 비트에서 다른 비트로 변화할 때 주어질 수 있는 소스의 확률을 버리는 결과를 가진다. 또한, Viterbi 복호기를 통해 나온 시퀀스가

H.263 복호기에서 제대로 복호를 못하는 코드를 가지게 된다.

이러한 가능성을 위의 두 제한을 이용하고 한 state에서 다른 state로 변화하는 소스의 전이 확률을 이용해 더 정확한 정보를 복원할 수 있었다.

하지만, 여기서 다른 소스의 전이 확률은 심볼들이 가지는 히프만 코드만을 이용하였다. 움직임 벡터나 DCT값들은 이전 프레임과 다음 프레임 혹은 이어있는 블록사이에 연관성이 존재하기 때문에 이러한 정보를 이용한다면 더 정확한 데이터를 복원할 수 있을 것이다.

### 참고문헌

- [1]K. Sayood and J. C. Borkenhagen, "Use of Residual Redundancy in the Design of Joint Source/Channel Coders", IEEE Trans. Comm., vol. 39, No. 6, June 1991.
- [2]N. Demir and K. Sayood, "Joint Source/Channel Coding for Variable Length Codes", DCC '98 Proceedings, 1998.
- [3]N. Phamdo and N. Farvardin, "Optimal Detection of Discrete Markov Sources Over Discrete Memoryless Channels - Applications to Combined Source-Channel Coding", Information Theory, 1993. Proceedings, 1993
- [4]M. Park and D. J. Miller, "Joint Source-Channel Decoding for Variable-Length Encoded Data by Exact and Approximate MAP Sequence Estimation", IEEE Trans. Comm., Vol. 48, No. 1, January 2000.
- [5]A. H. Murad and T. E. Fuja, "Exploiting the Residual Redundancy in Motion Estimation Vectors to Improve the Quality of Compressed Video Transmitted over noisy channels", ICIP 98. Proceedings. 1998.
- [6]C. E. Shannon, " A mathematical theory of communication", Bell Syst. Tech. J., Vol. 27, 1948.
- [7]A. J. Viterbi and J. K. Omura, "Principles of Digital Communication and Coding", McGraw-Hill, 1979.
- [8]Recommendation H.263: Video coding for low bit rate communication, ITU-T, September, 1997.