

## 신경회로망을 이용한 비디오의 고정 비트율 제어에 관한 연구

正會員 李 八 鎭\*, 姜 寅 坤\*, 具 顯 書\*, 朴 聖 宇\*\*, 金 永 川\*

### A Study on the Fixed Bit Rate Transmission of Video using Neural Networks

Pal Jin Lee\*, In Kon Kang\*, Hyun Seo Koo\*, Seong Woo Park\*\*, Young Chon Kim\* Regular Members

#### 要 約

본 논문에서는 압축된 디지털 비디오 소스를 일정한 비트율을 갖는 채널에 전송하기 위하여 신경회로망을 이용한 새로운 고정 비트율 제어 방법을 제안하였다. 제안된 고정 비트율을 제어기는 3층 구조의 완전 연결된 신경회로망으로 구성된다. 신경회로망의 구조는 16개의 입력 노드와 6개의 은닉 노드, 그리고 3개의 출력 노드를 갖는다. 가변 출력 비트율과 처리되는 영상 사이의 관계를 학습하기 위하여 Back Propagation 알고리즘을 이용하였다. 고정 비트율 제어를 위한 학습 데이터는 비트 분포 뿐만 아니라 부호화기에서 비트 발생률에 영향을 미치는 여러가지 인자로 구성하였다. 제안된 고정 비트율 제어기의 성능 평가를 위하여 BPP, PSNR, 양자화 파라미터의 분포 등을 측정하였으며, ISO MPEG의 TM1과도 그 결과를 비교하였다. 비교결과, 화질이나 대역 이용률 측면에서 신경회로망을 이용한 제어기법이 TM1보다 보다 더 좋아짐을 알 수 있었다. 시플레이션은 MPEG 표준영상인 Football, Tabletennis를 이용하여 4종류의 GOP를 구성하여 수행하였다.

#### ABSTRACT

In this paper, we propose a new fixed bit rate controller method using neural networks to improve the QOS in transmitting a compressed digital video source through a constant rate channel. The proposed fixed bit-rate controller proposed consists of the three-layered fully connected neural network with 16 input, 6 hidden, and 3 output nodes. Back propagation algorithm is used for learning of the neural network. The training data include various factors that influence the generation of bit rate in an encoder. To

\*전북대학교 컴퓨터공학과  
Dept. of Computer Engineering, ChonBuk  
National Univ.

\*\*한남대학교 정보통신공학과  
Dept. of Information and Communication Eng.,  
HanNam Univ.  
論文番號 : 94148  
接受日字 : 1994年 6月 8日

evaluate the performance of the proposed controller, we measure and evaluate the performance in terms of BPP, PSNR, and the distribution of the quantizer step size, compare to those of TM1 of ISO-MPEG. The results of simulation show that the proposed algorithm is superior to the conventional one for the picture quality and bandwidth utilization aspects. Simulation is performed for four types of a GOP with Football and Tabletennis which are MPEG standard images.

## I. 서 론

MPEG은 ISO-IEC/JTC1/SC29/WG11에 소속되어 디지털 저장 매체용 동영상 압축 기법을 연구하는 위원회이다. MPEG에서 권고하는 압축 기법에 따르면 동영상의 전송은 동영상의 특성을 고려하여 영상의 시 공간적인 중복성을 제거한 압축과 양자화 및 통계적인 특성에 따른 가변 길이 부호화 등의 일련의 압축과정을 수행하여 최소한의 정보로 영상 전송이 이루어진다. 그러나 부호화기에서 여러 압축 과정을 거쳐 생성되는 정보는 영상의 특성에 따라 다양하고 가변적이기 때문에 양질의 화질을 보장하기 위해서 할당된 대역을 효율적으로 이용할 수 있는 비트율 제어가 요구된다.

최근에 비트율 제어를 위한 다양한 방법이 국제 표준안을 비롯한 많은 문헌에 소개되고 있다.<sup>(1)(2)(3)</sup> 이에 따르면 양자화 파라미터의 크기를 조절하여 양자화 과정에서 발생하는 정보량을 조절함으로써 비트율 제어를 시도하고 있다. 그러나 양자화는 정보 손실을 야기하는 압축이기 때문에 제한된 채널 활용을 극대화하여 양질의 화질을 전송할 수 있도록 하고, 버퍼에서는 오버플로우(overflow)나 언더플로우(underflow)가 발생하지 않도록 하기 위해서는 적절한 양자화 파라미터를 결정하는 것이 매우 중요한 문제이다. SM3(Simulation Model 3)나 TM1(Test Model 1)에서는 양자화 파라미터를 결정하기 위하여 버퍼의 상태나 3단계를 이용한 비트율 제어 방법을 채택하고 있으나 불규칙하게 변하는 영상의 특성을 효율적으로 반영하지 못하기 때문에 복원 영상의 균일한 화질을 유지하기 어렵다.

본 연구에서는 가변적이고 예측하기 힘든 동영상 압축 데이터를 효율적으로 전송하기 위하여 영상의 특성 뿐만 아니라 이에 따른 버퍼의 상태를 파악하여 양자화 단계 값을 적응적으로 조정함으로써 원 영상에 대한 복원 영상의 균일한 화질 유지 및 효율적인 채널사용이 가능하도록 신경회로망을 이용한 새로운 고정 비트율 제어 방

법을 제안하였다.

이를 위하여 신경회로망의 입력 노드 16개중 10개의 노드에는 현재까지의 비트 발생 경향, 나머지 6개에는 비트 발생률에 영향을 미치는 매크로블럭의 유형, 매크로블럭의 평균 화소값, 매크로블럭의 활동도, 매크로블럭의 목표 비트, 실제 버퍼의 총만도, 이전 매크로블럭의 양자화 파라미터를 사용하였고, 출력으로는 양자화 단계값의 증가, 유지, 감소를 이용하였다.

제안된 고정 비트율 제어기의 성능 평가 및 분석은 영상군(GOP)을 구성하는 프레임의 수를 나타내는 N과 B 프레임의 수를 나타내는 M에 따라 4종류의 GOP 유형을 구성하여 수행하였다.

## II. MPEG 비디오의 고정 비트율 제어

MPEG 비디오에서는 고정 대역을 통한 효율적인 동영상의 전송을 위해서 동영상의 다양한 특성을 이용하여 여러가지 압축을 실시한다. 그러나 압축과정을 거쳐 부호화기에서 발생하는 압축 정보는 VBR(Variable Bit Rate) 특성을 가지고 있기 때문에 제한된 용량의 통신 채널을 효율적으로 이용하기 위해서는 부호화기와 채널 사이에 버퍼를 이용하여 일정량의 정보를 채널로 전송할 수 있어야 한다.

ISO-IEC/JTC1/SC29/WG8의 MPEG-1 급 SM3에서는 비트율 조절을 위한 버퍼 제어가 양자화 파라미터만으로 가능하다고 정의하고 있다.<sup>(3)</sup> 영상의 3가지 유형 I, P, B를 이용하여 균일한 화질을 유지하기 위해서는 각각의 유형에 적당한 비트 할당이 요구되며, 전송되는 영상의 특성에 따라 그 분포는 다양하게 나타난다. 이 때, 양자화 파라미터의 조절은 버퍼 상태(총만도)만을 반영하기 때문에 선형적으로 이루어진다.

TM1에서는 피드백되는 버퍼 총만도 뿐만 아니라, 처리될 영상에 할당된 비트수, 가상 버퍼 총만도, 영상의 복잡도 및 매크로블럭의 활동도 등을 함께 고려한 즉,

이전의 정보와 예측 정보를 이용한 적응적인 양자화 파라미터의 결정으로 비트 발생을 제어하고 있다.<sup>(11)</sup> TM1에서의 비트율 제어를 위한 양자화 단계값의 결정 과정은 프레임 단위의 목표 비트 할당, 가상 버퍼 충만도에 의한 기준 양자화 파라미터의 설정, 영상내에서의 공간 활동도를 이용한 적응적인 양자화 파라미터 결정 등의 3단계로 이루어진다.

고정된 전송 용량내에서 정보의 손실로 인한 화질 저하를 최소로 하면서 효율적인 영상 전송이 이루어질 수 있도록 하기 위해서는 각 영상 또는 각 블록의 특성과 인간의 시각 특성을 고려한 적응적인 비트 할당이 이루어져야 하는데 이것은 채널의 전송 용량과 이전 영상의 통계적 특성을 고려하는 것이 효율적이다. 또한, 동영상의 다양한 특성은 불규칙한 비트 발생을 유발하며 버퍼의 문제를 야기하기 때문에 버퍼의 충만도와 처리될 영상의 특성을 함께 고려한 적응적인 양자화 파라미터의 조절로 비트율을 제어해야 한다. 이와 같은 제어는 불규칙하게 발생하는 부호화기의 비트량을 일정하게 유지할 수 있을 뿐만 아니라 채널을 효율적으로 이용할 수 있기 때문에 화질을 균일하게 유지할 수 있다.

### Ⅲ. 신경회로망을 이용한 비트율 제어

#### 3.1 비트율 제어를 위한 신경회로망의 구조

비트율 제어를 위해 본 연구에서 제안한 신경회로망의 구조는 3층 구조의 Fully-Connected Neural Network으로서 그림 3.1과 같은 개념적 구조 모델을 이용하였다. 신경회로망을 이용한 학습은 Back Propagation 알고리즘을 이용하였다. 이 알고리즘은 Delta Rule을 이용하여 노드간의 결합강도를 수정하므로써 실제 출력과 기대 출력사이에서 발생하는 오차를 줄여 주어진 입력 패턴에 대해 올바른 학습을 수행하도록 하는 지도 학습 알고리즘이다.<sup>(10)(11)</sup>

Back Propagation 알고리즘의 학습 과정은 식(3.1)과 같이 입력 벡터에 대한 망의 출력으로 기대 출력과의 오차를 구하고, 이 오차에 따라 결합강도를 변화시킬 방향(+, -)과 변경치를 결정한 다음, 식(3.2)를 이용하여 현재의 결합강도를 변화시킨다.

$$\delta_{pk} = (y_{pk} - o_{pk})f'_k(\text{net}_{pk}^o) \quad (3.1)$$

여기서  $y_{pk}$ 는 기대 출력값을,  $o_{pk}$ 는 실제 출력값을 나타낸다.

$$W_{ki}(t+1) = W_{ki}(t) + \Delta_p W_{ki}(t) \quad (3.2)$$

여기서

$$\Delta_p W_{pk} = \eta (y_{pk} - o_{pk})f'_k(\text{net}_{pk}^o) i_{pi} \quad (3.3)$$

이다. 모든 학습 패턴의 기대값이 출력될때까지 결합강도의 수정과정은 반복된다. 이 때, 각 계층의 출력을 결정하기 위해서 사용되는 활성화 함수는 식(3.4)와 같은 미분 가능한 비선형 Sigmoid 함수를 이용한다.

$$f'_k(\text{net}_{pk}^o) = (1 + e^{-\text{net}_{pk}^o})^{-1} \quad (3.4)$$

입력층, 은닉층, 그리고 출력층의 각각에 대한 노드 수의 결정은 학습 시간, 비트율 제어를 통한 화질의 객관적인 평가 및 이에 따른 대역의 이용률을 고려하였다. 그리고 효율적인 비트율 제어를 수행하기 위해 다양한 시뮬레이션을 통하여 비트 발생량에 영향을 미치는 적합한 제어인자를 추출하였다.

#### 3.1.1 입력 노드의 구성

본 연구에서는 양자화 단계값을 결정하는데 버퍼의 충만도에 매크로블럭의 특성을 정확하게 반영시키므로써 효율적인 비트율 제어를 수행하여 양자화 단계값이 적응적으로 결정될 수 있도록 하였다. 따라서, MPEG 비디오의 일련의 압축과정을 거치는 동안 다양한 실험을 통하여 비트 발생의 원인이 되는 요소들과 버퍼의 상태 및 이전의 비트 발생 경향을 표 3.1과 같이 신경회로망의 입력값으로 결정하였다.

##### ① 매크로블럭의 유형

매크로블럭의 유형은 주로 P와 B를 처리하는 과정에서 정의된다. Interframe 부호화는 움직임 추정 및 보간에 의하여 수행되는데 먼저 움직임 보상 결정을 수행한다. 즉, 움직임 벡터값을 디코더로 보낼것인지 아닌지를 결정하게 되는데, 움직임 벡터값이 0이거나 또는 움직임 보상으로 인하여 에러가 커지는 경우에는 에러의 절대값에 따라 intra/Non-intra 부호화를 결정

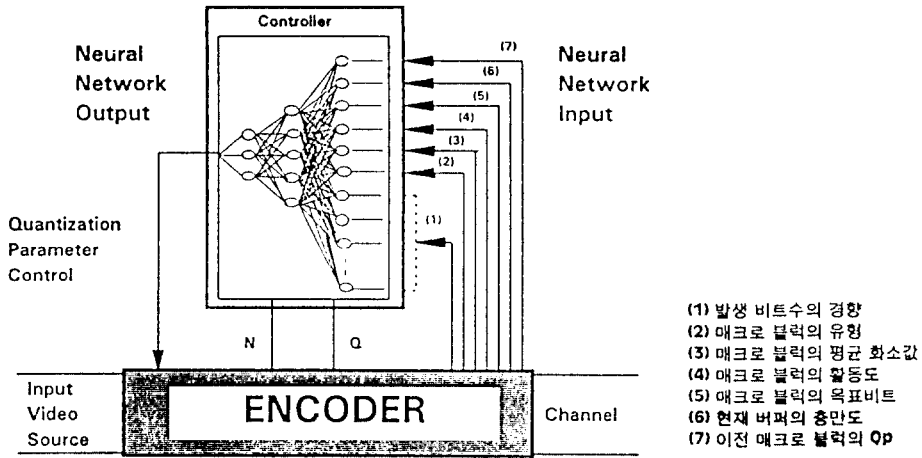


그림 3.1 신경회로망을 이용한 고정 비트율 제어기  
 Fig. 3.1 Fixed bit rate controller using neural networks

하고, 다음으로 부호화의 여부(coded/Not-coded)를 결정한다. 이와 같은 매크로블럭의 유형의 결정은 비트 발생률에 커다란 영향을 미치게 되는데, 이는 영상의 복잡도가 증가할수록 발생 비트량도 증가하기 때문이다.

② 매크로블럭의 평균 화소값

매크로블럭의 평균 화소값과 화소의 분포는 비트 발생에 많은 영향을 주는 요소이다. 양자화 과정을 수행하면 양자화 손실이 발생하기 때문에 해당 블럭에 대한 정확한 비트 추정어 어렵다. 이를 보완하는 하는 것으로 매크로블럭을 구성하고 있는 화소값의 분포는 DCT 및 양자화에 많은 영향을 주기 때문에 식(3.5)와 같이 화소값의 평균( $M_p$ )을 하나의 입력으로 선택하였다.

$$M_p = \frac{1}{64} \sum_{k=1}^{64} (P_k) \quad (3.5)$$

여기서  $P_k$ 는  $8 \times 8$  블럭을 구성하는 화소값을 나타낸다.

③ 매크로블럭의 활동도

공간 활동도는 매크로블럭내의 휘도 성분 블럭을 구성하는 화소의 분산에 의해 구해지며, 이것을 정규화한 값과 구해진 기준 양자화 파라미터를 곱하여 적용적 양자

화 단계값(mquant)을 결정한다. j번째 매크로블럭을 구성하는 4개의 휘도 성분 블럭으로부터 공간 활동도를 구하는 식(3.6)과 같다.

$$act_j = 1 + \min(var\_sblk) \quad (3.6)$$

$sblk = 1, 4$

$$var\_sblk = \sum_{k=1}^{64} (P_k - M_p)^2 \quad (3.7)$$

$P_k$ 는 화소값,  $M_p$ 은 한 블럭의 화소의 평균을 나타낸다.

④ 매크로블럭의 목표 비트

목표 비트의 할당은 영상 단위로 이루어지며, 영상의 복잡도와 GOP를 구성하는 각 유형의 영상수( $N_b$ ,  $N_p$ )를 이용하여 구한다. 영상의 복잡도는 영상의 유형마다 이전 영상의 양자화 단계값의 평균과 이전 영상을 처리하며 발생한 비트수를 사용하여 구한다. 식 (3.8), 식 (3.9), 식 (3.10)은 영상 유형에 따른 목표 비트 계산식을 나타낸다. ( $K_b$ ,  $K_p$ 는 'universal' 상수)

식 (3.8), 식 (3.9), 식 (3.10)에 의해 할당된 영상 단위의 목표 비트를 식 (3.11)과 같이 하나의 영상을 구성하고 있는 매크로블럭수(MB\_cnt)로 나눈 값이다.

$$T_i = \max \left\{ \frac{R}{1 + \frac{N_p * X_p}{X_i * K_p} + \frac{N_b * X_b}{X_i * K_b}}, \right. \\ \left. bit\_rate / (8 * picture\_rate) \right\} \quad (3.8)$$

$$T_p = \max \left\{ \frac{R}{N_p + \frac{N_b * K_b * X_b}{K_b * X_p}}, \right. \\ \left. bit\_rate / (8 * picture\_rate) \right\} \quad (3.9)$$

$$T_b = \max \left\{ \frac{R}{N_b + \frac{N_p * K_p * X_p}{K_p * X_b}}, \right. \\ \left. bit\_rate / (8 * picture\_rate) \right\} \quad (3.10)$$

$$(MT_i, MT_b, MT_p) = (T_i, T_b, T_p) / MB\_cnt \quad (3.11)$$

⑤ 현재의 실제 버퍼의 충만도

TM1에서는 3가지 유형의 영상 각각에 가상 버퍼를 두고 매크로블럭 단위로 영상에 할당된 비트수, 초기의 버퍼 충만도, 그리고 실제 발생한 비트수를 이용하여 가상 버퍼의 충만도를 식(3.12)에 의해 (d<sub>t</sub>)를 구하고, 이를 이용하여 기준 양자화 파라미터를 결정한다.

$$d'_t = d'_o + B_{j-1} - T_t \frac{(j-1)}{MB\_cnt} \quad (3.12)$$

여기서 t는 각 프레임의 유형(I, P, B)을 나타낸다. d<sub>o</sub>'는 각 유형에 대한 버퍼의 초기상태를 나타내며, d<sub>t</sub>'는 가상버퍼의 충만도를 나타낸다.

매크로블럭의 특성(영상의 특성)은 버퍼의 충만도에 많은 영향을 미친다. 일정한 비트율로 데이터를 전송하기 위하여 버퍼의 현재 상태를 입력값으로 사용하였다.

⑥ 이전 매크로블럭의 양자화 파라미터

비트율 제어를 위한 여러가지 제어인자를 이용한 양자화 파라미터의 결정은 이전의 양자화 파라미터를 기준으로 결정된다. 따라서, 비트량의 제어를 위해 요구되는 양자화 파라미터를 결정하기 위하여 이전의 양자

화 파라미터를 하나의 입력으로 선택하였다.

⑦ 현재까지의 비트 발생 경향

동영상의 압축은 intraframe 과 interframe 부호화를 이용하여 수행된다. 비디오 시퀀스를 구성하는 각 GOP 내에서 인접한 영상 또는 인접한 슬라이스나 매크로블럭들 사이에는 상관성이 존재하기 때문에 서로 비슷한 비트 발생률을 갖는다. 이러한 영상의 특성은 목표 비트량에 영향을 미치기 때문에 현재까지의 비트 발생 경향을 입력으로 사용하였다.

표 3.1 입력 패턴의 구성  
Table 3.1 Construction of input patterns

입력내용	할당된 노드수
매크로블럭의 유형	1
매크로블럭의 평균 화소값	1
매크로블럭의 활동도	1
매크로블럭의 목표 비트	1
현재 버퍼의 충만도	1
이전 매크로블럭의 QP	1
33개 매크로블럭의 비트수	10
합 계	16

3.1.2 출력 노드의 구성

본 연구에서 제안한 비트율 제어기의 출력으로는 균일한 화질 유지를 목적으로 인접 매크로블럭간 양자화 파라미터의 차이를 one step으로 고정시켜 이전의 양자화 파라미터를 기준으로 입력 패턴에 따라 현재의 양자화 파라미터를 증가, 감소, 유지로 표현하는 값을 출력하도록 3개의 노드를 운영하였다. 출력값에 따라 각 노드가 갖는 상태값들은 표 3.2와 같다.

표 3.2 신경회로망의 출력값  
Table 3.2 Output node values of neural network

제어 \ 노드	출력 1	출력 2	출력 3
증 가	1	0	0
유 지	0	1	0
감 소	0	0	1

### 3.2 고정 비트율을 위한 데이터의 학습

신경회로망의 학습이란 시스템의 목적에 맞도록 노드 간의 결합강도를 조정하는 것으로, 입력으로 가능한 모든 패턴에 대해서 만족하는 결과를 출력하는 신경회로망의 결합강도를 찾는 과정은 학습용 입력 패턴의 수에 따라 요구되는 시간이 다양하다. 올바른 비트율 제어를 위한 안정된 최적의 결합강도를 얻기 위해서는 수많은 반복 학습이 요구된다. 그러나 입력으로 가능한 모든 영상을 이용하여 학습을 실시하는 것은 불가능하기 때문에 본 연구에서는 영상에 대한 양자화 파라미터를 기준으로 비트율의 증가, 유지, 감소 3가지 클래스로 분류하여 학습하였다. 클래스는 영상의 GOP를 이용하여 매크로블럭당 발생하는 비트수를 기준으로 이전의 매크로블럭 비트수와 비교하여 분류하였다.

제한한 신경회로망을 이용한 비트율 제어기의 학습은 Football 영상에 대해 휘도성분만을 가지고  $N=15$ ,  $M=2$ 인 시퀀스를 처리하여 발생한 비트중에서 영상군에 해당하는 비트 발생 경향과 비트율을 효율적으로 제어하는 6가지 제어인자를 추출하여 학습용 입력 패턴으로 사용하였다. 또한 양자화 파라미터의 효율적인 제어를 위하여 요구되는 기대값의 학습 데이터는 발생한 매크로블럭당 비트수의 증가, 유지, 감소의 3가지로 분류하여 적용하였다. 학습 데이터를 이용한 학습 방법은 신경회로망의 입력으로 사용되는 16개의 노드 중에서 10개에는 한 프레임 단위를 입력한 동시에 각 노드에는 매크로블럭 단위로 발생한 비트 33개를 스위치하면서 학습되도록 하였고, 나머지 6개의 노드에는 가변적이고 예측하기 힘든 영상의 특성에 따라 발생하는 비트 분포에 영향을 미치는 제어인자를 입력으로 하였다. 처음 10개의 노드에는 비트수가 그룹단위로 형성되어 입력되고, 나머지 6개에는 각각 제어인자가 1:1로 할당되어 입력된다.

학습을 위한 초기환경은 다음과 같다. 초기에 주어지는 유니트간의 결합강도와 바이어스는 난수 발생기를 이용하여 -0.5에서 0.5 범위내에서 부여하였고, 학습 과정에 많은 영향을 주는 learning rate parameter( $\eta$ )는 0.001에서 0.01 범위내에서 설정하였다.  $\eta$ 는 학습의 수렴속도를 조절하는 파라미터로서 설정된 초기치에 의해 이 값이 크면 수렴속도가 빠르나 local minima에 도달할 염려가 있다. 또한 이 값이 작을 경우 수렴속도가 너무 늦어 실시간 제어시 문제가 된다. 본 연구에

서는 이와 같은 문제점을 피하기 위하여 학습 테이블의 추출값이 초기화 자료일때와 보완된 값일때를 구분하여 적용하였다. 초기화 값에 대한 학습은 정확한 값이 아니므로 낮은 학습률을 적용하였고, 보완된 값에 대해서는 수렴속도를 감안하여 높은 학습률을 적용하여 최적의 결합강도에 수렴할 수 있도록 하였다.

## IV. 시뮬레이션 및 성능평가

고정 비트율 영상 전송 시스템에 신경회로망을 적용하여 비트율 제어를 하기 위해서 MPEG 비디오 압축 기법을 기반으로 영상 전송 시스템을 SUN SPARC Station에서 X-Window와 C 언어를 사용하여 구현하였다. 시뮬레이션에 사용된 영상은 MPEG의 표준영상들 중 Football, Tabletennis 두 가지 영상이며, 영상포맷인 SIF의 휘도 성분만을 추출한 영상으로 구성하였다.

### 4.1 시뮬레이션

시뮬레이션을 위한 동영상 전송 시스템의 초기 조건은 통신채널의 전송용량을 1.5 Mbps로, 초당 전송되는 영상 수를 30개로 설정하였다. 영상 시퀀스는 두 영상에 대하여 각각 5초 분량인 150개 영상을 사용하였다.

제한된 전송 용량내에서 적응적인 비트율 제어를 위해서는 GOP를 구성하고 있는 각 유형의 영상수에 따라 화질과 압축률을 동시에 만족시킬 수 있는 비트 할당이 이루어져야 한다. 비트 할당은 GOP 단위로 각 유형의 영상에 순차적으로 할당하는데, 화질을 제어하는 I와 P 영상에 상대적으로 많은 비트가 할당되고, 반면에 압축효과를 높이기 위해 B영상에는 적은 비트가 할당된다.

시뮬레이션은 영상 시퀀스의 구성형태를 나타내는 GOP의 길이와 구성을 나타내는 N, M을 각 유형의 영상에 할당되는 비트수의 변화에 따라 표4.1과 같이 시퀀스 I, II, III, IV로 분류하여 실시하였다.

또한, 본 연구에서 제안한 비트율 제어기의 성능 평가를 수행하기 위해서 화질은 영상에 대한 식 (4.1)의 PSNR(Peak Signal to Noise Ratio)과 매크로블럭에 대한 PSNR의 표준편차를 사용하였고, 비트율은 영상 단위로 BPP(Bits Per Picture)를 사용하였다. 그리고 화질과 비트율에 큰 영향을 미치는 양자화 파라메

표 4.1 각 유형별 영상군의 구성  
Table 4.1 Construction of GOP per each type

시퀀스	N	M
I	1	0
II	10	2
III	10	3
IV	15	3

터(QP)의 분포도 비교하였다.

$$PSNR = 20 \log_{10}(255/RMSE) \quad (4.1)$$

#### 4.2 성능평가 및 분석

신경회로망을 이용한 비트율 제어기의 성능평가는 평가인자를 이용하여 TM1의 비트율 제어 기법과 시퀀스 별로 비교하였다.

##### 4.2.1 시퀀스 I의 성능평가

신경회로망을 이용한 비트율 제어기에 학습용 입력 패턴의 수를 변화시키며 실시한 학습 결과를 적용하여 즉, 적용적인 비트율 제어에 신경회로망 응용의 가능성 여부를 알아보기 위하여 임의의 학습패턴을 30개, 150개, 3000개로 증가시켜 학습을 실시하고 그 결과를 비트율 제어기에 적용하였고, 시퀀스 I의 영상 시퀀스를 사용하여 TM1과 학습 패턴 수의 변화에 따른 결과를 비교하였다. 표4.2는 학습패턴을 3000개로 했을 때 두 영상의 평균 결과이고, 그림 4.1, 그림 4.2는 각각 Football의 PSNR과 BPP이다.

표 4.2 시퀀스 I의 평균 결과  
Table 4.2 Average results for the sequence type I

평가인자 방법	Football		Tabletennis	
	PSNR	BPP	PSNR	BPP
TM1	34.359	138686.1	25.943	138749.4
BPN30	30.743	75502.7	23.445	84494.7
BPN150	32.816	143061.2	25.183	135968.5
BPN3000	34.860	138653.1	26.135	138663.4

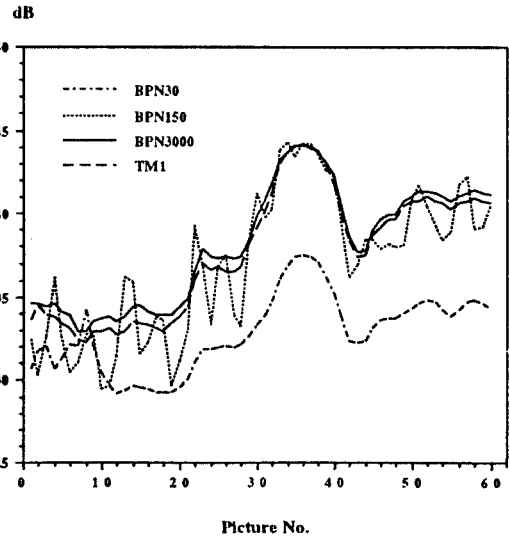


그림 4.1 시퀀스 I에 대한 Football의 PSNR  
Fig. 4.1 PSNR of Football to the sequence type I

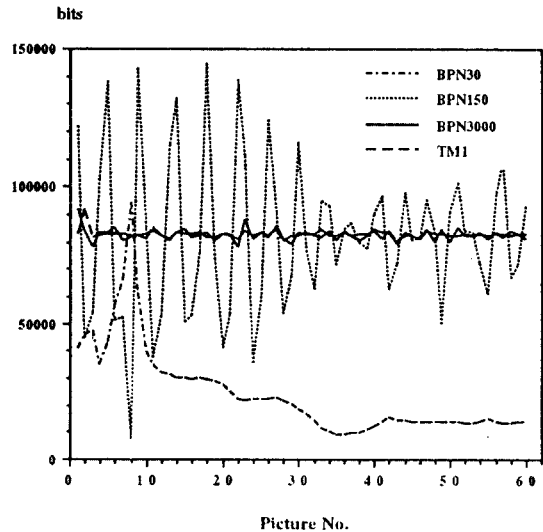


그림 4.2 시퀀스 I에 대한 Football의 BPP  
Fig. 4.2 BPP of Football to the sequence type I

그림 4.1, 그림 4.2에 나타난 결과는 신경회로망이 처음 30개 패턴으로 학습되었을 때, 다음에 150개의 패턴으로 학습되었을 때 결과는 점진적으로 향상됨을 볼 수 있고, 3300개의 패턴으로 학습된 결과를 이용했을 때는 TM1보다 훨씬 나아짐을 볼 수 있다. 이는 학

습된 패턴의 수가 많아지면 많아질수록 부호화기에서 불규칙하게 발생하는 정보량을 효율적으로 제어할 수 있고, 동시에 화질이 향상됨을 볼 수 있다. 결과적으로, 부호화기의 비트율 제어에 영향을 미치는 여러가지 요소들을 결합하여 만들 수 있는 많은 입력패턴의 학습이 가능한 경우 적응적인 비트율 제어에 신경회로망의 응용이 적합함을 말해주고 있다. 여기서 BPN은 신경회로망을 이용한 제어기법을 나타낸다.

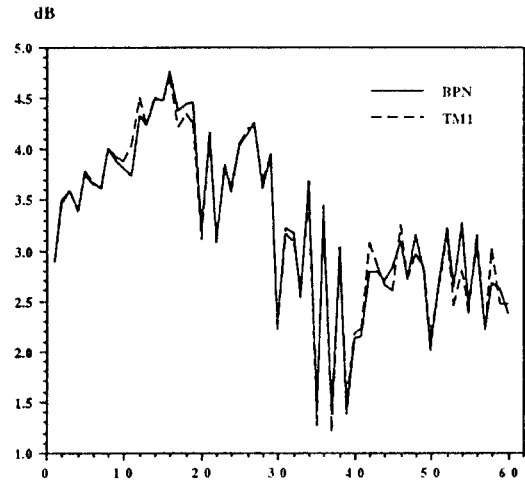
4.2.2 시퀀스 II의 성능평가

N=10이고 M=2인 영상 시퀀스 II는 초당 전송되는 GOP가 3개이며, 하나의 GOP는 각각 I 영상이 1개, P 영상이 4개, B 영상이 5개로 이루어져 있다. 따라서 하나의 GOP에 할당되는 비트량은 0.5 Mbps(1.5 Mbps/3)이다. 표4.3, 그림4.3, 그림4.4는 시퀀스 II를 사용하여 시뮬레이션을 수행한 결과이다.

시퀀스 II의 실험 결과는 비트 발생과 복원 화질에 큰 영향을 미치는 양자화 파라미터와 비트 발생량은 Football의 경우 신경회로망을 이용하여 제어한 결과가 작았고, Tabletennis 경우는 반대로 나타났다. 반면에 전체적인 복원 화질은 두 영상 모두 신경회로망을 이용한 것이 더 나았음을 알 수 있고, 또한 매크로블럭의 PSNR 표준편차는 TM1보다 신경회로망을 이용한 경우가 작게 나타남으로써 인접한 매크로블럭 사이의 화질의 변화가 심하지 않아 전체적인 복원 영상의 화질은 신경회로망을 이용한 것이 더 나았음을 알 수 있다.

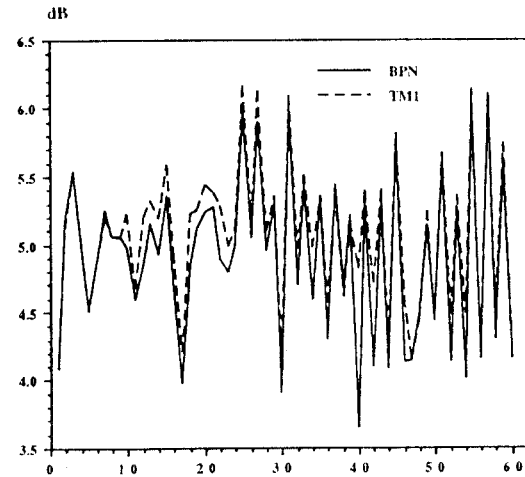
표 4.3 시퀀스 II의 평균 결과  
Table 4.3 Average results for the sequence type II

평가인자	Football		Tabletennis	
	TM1	BPN	TM1	BPN
BPP	52722.3	52706.9	53066.7	53125.8
QP	3.093	3.074	4.315	4.334
PSNR	36.126	36.134	30.656	30.692
PSNR의 표준편차	3.352	3.291	5.140	4.918



Picture No.

그림 4.3 시퀀스 II에 대한 Football의 PSNR 표준편차  
Fig. 4.3 Stand. Dev. of Football for the sequence type II



Picture No.

그림 4.4. 시퀀스 II에 대한 Tabletennis의 PSNR 표준편차  
Fig. 4.4 Stand. Dev. of Tabletennis for the sequence type II

4.2.3 시퀀스 III의 성능평가

N=10이고 M=3인 영상 시퀀스 III는 초당 전송되는 GOP가 3개이며, 하나의 GOP는 각각 I 영상이 1개, P 영상이 3개, B 영상이 6개로 이루어져 있다. 따라서 하나의 GOP에 할당되는 비트량은 0.5 Mbps(1.5



Mbps/3)이다. 표 4.4, 그림 4.5, 그림 4.6은 시퀀스 III를 사용하여 시뮬레이션을 수행한 결과이다.

시퀀스 III의 실험 결과는 시퀀스 II와 마찬가지로 Football를 처리한 결과는 복원화질에 큰 영향을 미치는 양자화 파라미터와 비트 발생량은 신경회로망을 이용하여 제어한 결과가 작았고, Tabletennis 경우는 반대로 나타났다. 그러나 전체적인 복원화질은 두 영상 모두 신경회로망을 이용한 것이 더 나았음을 알 수 있고, 또한 매크로블럭의 PSNR 표준편차는 TM1보다 신경회로망을 이용한 경우가 작게 나타남으로써 인접한 매크로블럭 사이의 화질의 변화가 심하지 않아 전체적으로 복원된 영상의 화질은 신경회로망을 이용한 것이 더 나았음을 알 수 있다.

표 4.4 시퀀스 III 평균 결과  
Table 4.4 Average results for the sequence type III

평가인자	Football		Tabletennis	
	TM1	BPN	TM1	BPN
BPP	52431.1	52406.0	52421.1	52442.8
QP	3.171	3.138	4.422	4.445
PSNR	36.112	36.150	30.508	30.468
PSNR의 표준편차	3.260	3.259	5.132	5.000

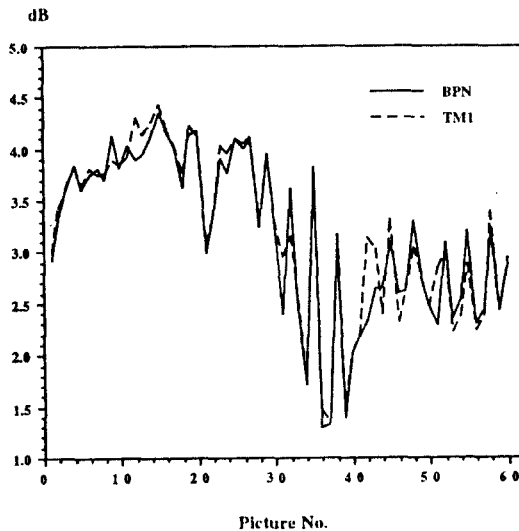


그림 4.5 시퀀스 III에 대한 Football의 PSNR 표준편차  
Fig. 4.5 Stand. Dev. of Football for the sequence type III

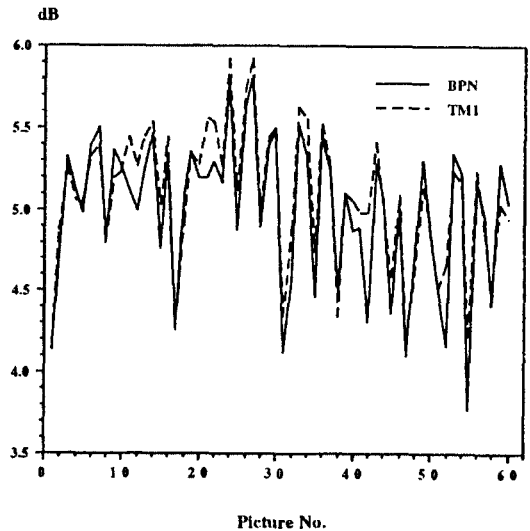


그림 4.6 시퀀스 III에 대한 Tabletennis의 PSNR 표준편차  
Fig.4.6 Stand. Dev. of Tabletennis for the sequence type

#### 4.2.4 시퀀스 IV의 성능평가

N=15이고 M=3인 영상 시퀀스 IV는 초당 전송되는 GOP가 2개이며, 하나의 GOP는 각각 I 영상이 1개, P 영상이 4개, B 영상이 10개로 이루어져 있다. 따라서 하나의 GOP에 할당되는 비트량은 0.75 Mbps(1.5 Mbps/2)이다. 표 4.5, 그림 4.7, 그림 4.8은 시퀀스 IV를 사용하여 시뮬레이션을 수행한 결과이다.

영상 시퀀스 IV를 사용한 시뮬레이션 결과는 Football과 Tabletennis 두 영상에 대하여 성능 평가 인자인 비트 발생량, 양자화 파라미터, PSNR이 모두 신경회로망을 이용한 제어의 결과가 더 좋게 나타났다. 신경회로망을 이용한 비트율 제어 결과, 양자화 파라미터가 작으며 비트 발생량이 적고, 비트 발생량이 적으면서도 PSNR이 높게 나타난 것은 영상을 구성하고 있

표 4.5 시퀀스 IV 평균 결과  
Table 4.5 Average results for the sequence type IV

평가인자	Football		Tabletennis	
	TM1	BPN	TM1	BPN
BPP	53046.3	52812.5	53176.9	53113.5
QP	3.103	3.031	4.169	4.147
PSNR	36.070	36.116	30.745	30.799
PSNR의 표준편차	3.283	3.128	5.038	4.764

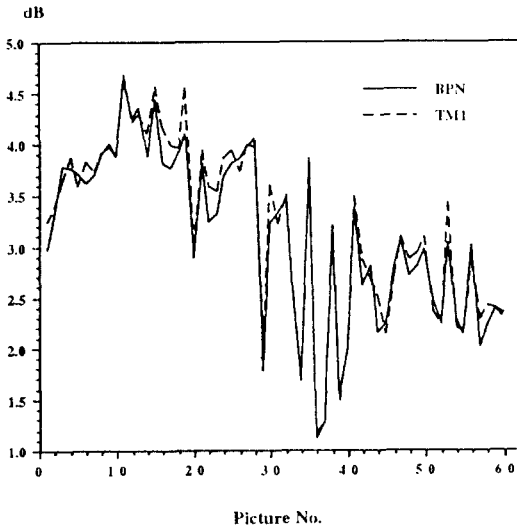


그림 4.7 시퀀스 IV에 대한 Football의 PSNR 표준편차  
Fig. 4.7 Stand. Dev. of Football for the sequence type

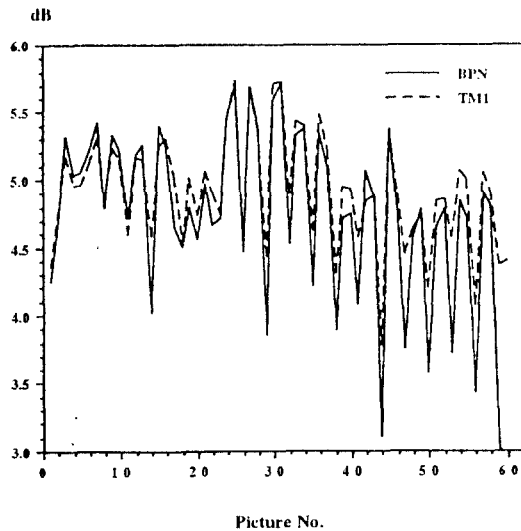


그림 4.8 시퀀스 IV에 대한 Tabletennis의 PSNR 표준편차  
Fig. 4.8 Stand. Dev. of Tabletennis for the sequence type

는 매크로블럭 사이의 화질의 변화가 심하지 않아 전체적으로 복원된 영상의 화질을 균일하게 유지할 수 있기 때문이다. 결과적으로, 제한된 전송 용량 내에서 보다 적은 비트로 양질의 화질을 유지하기 위한 비트율 제어에 신경회로망을 이용한 것이 성능이 우수함을 알 수

있다.

## V. 결 론

본 연구에서는 원 영상에 대한 복원 영상의 균일한 화질 유지와 채널의 사용을 극대화 하기 위한 효율적인 비트율 제어에 관한 연구를 수행하였다. 이를 위하여 복잡하고 가변적이며 예측하기 힘든 비선형적인 특성을 갖는 부호화기의 비트 발생률 제어를 위한 양자화 파라미터의 결정에 신경회로망의 학습기능을 이용하여 효율적이고 적응적인 제어 알고리즘의 구현 및 이에 대한 성능 평가를 실시하여 기존의 비트율 제어 방법과 결과를 비교하였다.

고정된 용량의 채널을 통한 영상 전송에서는 영상 시퀀스를 구성하고 있는 GOP의 구성 형태에 따라 각 유형 영상에 할당되는 비트량은 다양하고 또한, 복원된 화질에도 많은 영향을 미친다. 본 논문에서는 GOP를 여러가지 형태로 분류하여 시뮬레이션을 실시하였다. 그결과, 시퀀스 II 보다는 시퀀스 III가, 시퀀스 III 보다는 시퀀스 IV가 평균적으로 각 유형의 영상에 할당되는 비트수가 많아 더 나은 화질을 유지할 수 있었다.

동영상 전송 시스템의 비트율 제어에 신경회로망을 이용한 경우 비트율 면에서는 TM1과 유사하면서도 화질면에서는 전체 매크로블럭별 평균에 대한 표준편차가 TM1과 비교하여 적게 나타남으로써 인접한 매크로블럭 간의 화질 차이가 TM1보다 적어 전체적으로 균일한 화질을 유지할 수 있었다. 결과적으로, 부호화기에서 영상을 처리하는 동안 비트 발생에 영향을 미치는 원인들과 주기적인 유사성을 감안한 비트 발생의 경향을 학습 시킨 비트율 제어기가 TM1보다 화질면에서 우수함을 알 수 있었다.

그러나 Football과 Tabletennis 영상의 결과에서 다소 차이를 보는 원인은 개개의 영상이 내재하고 있는 특성 때문이다. 그러므로 신경회로망을 이용하여 고정 대역을 통한 영상 전송을 효율적으로 수행하기 위해서는 특정 영상에 국한되지 않는 다양한 학습 패턴의 구성과 비트 발생 분포에 영향을 미치는 제어인자의 선택이 매우 중요하다. 따라서, 많은 영상을 사용하여 보다 더 다양한 제어 인자의 도출과 학습 패턴의 분류 및 충분한 학습시간을 갖는다면 더 나은 제어 효과를 기대할 수 있을 것으로 생각한다.

참고문헌

1. International Organization for Standardisation, ISO-IEC/JTC1/SC29/WG11, Test Model 1, Draft Revision 1, May, 1992.
2. International Organization for Standardisation : Coded Representation of Pictures and Audio Information, ISO-IEC/JTC1/SC29/WG11 MPEG 92/160.
3. ISO-IEC/JTC1/SC29/WG8 MPEG90, MPEG Video Simulation Model Three(SM3), 1990.
4. Dider Le Gall, "MPEG : A Video Compression Standard for Multimedia Applications," ACM Comm., Vol.34, No.4, pp.46-58, April, 1991.
5. J.Pandel, "Variable bit-rate image sequence coding with adaptive quantization," Signal Processing : Image Communication 3, pp.123-128, 1991.
6. Tanaka, H.Haseyama, K.Manabe and F.Kishino, "Bit-rate control algorithm on video packet coding and its characteristics," Signal Processing : Image Communication 3, pp.321-332, 1991.
7. J.Katto, K.Onda and Y.Yasuda, "Variable bit-rate coding based on human visual system," Signal Processing : Image Communication 3, pp.313-320, 1991.
8. Richard C.Lau Paul E.Fleischer, Shaw-Min Lei, "Receiver Buffer Control for Variable Bit-Rate Real-Time Video," IEEE ICC '92, pp.544-550, 1992.
9. Eric Viscito, Cesar Gonzales, "A Video Compression Algorithm with Adaptive Allocation and Quantization," SPIE Vol. 1605 Visual Comm. and Image Processing '91 : Visual Comm., pp.58-72, Nov. 1991.
10. Patric K. Simpson, Artificial Neural System, Pergamon Press.
11. V.Vemuri, "Artificial Neural Networks : An Introduction," IEEE Computer Society Press, pp.1-12, 1988.
12. Richard P.Lippman, "An Introduction to Computing with Neural Nets," IEEE ASSP MAGAZINE, pp.4-22, 1987.



李 八 鎭 (Pal Jin Lee)      정희원  
 1964년 1월 15일생  
 1982. 3. ~1986. 2. : 조선대학교 공과대학 전산기공학과 (공학사)  
 1986. 9. ~1988. 8. : 중앙대학교 공과대학 컴퓨터공학과 (공학석사)  
 1990. 4. ~1992. 2. : 목포대학교 전산통계학과(조교)  
 1992. 2. ~1995. 현재 : 전북대학교 공과대학 컴퓨터공학과 (박사과정수료)

姜 寅 坤 (In Kon Kang)      정희원  
 1966년 4월 22일생  
 1992. 3. ~1995. 현재 : 전북대학교 공과대학 컴퓨터공학과 (박사과정수료)  
 제18권 12호 참조



具顯晝 (Hyun Seo Koo) 정회원

1967년 9월 3일생

1985. 3.~1992. 2. : 전북대학교 공  
과대학 컴퓨터공학과 (공  
학사)

1992. 3.~1994. 2. : 전북대학교 공  
과대학 컴퓨터공학과 (공  
학석사)

1994. 2~1995. 현재 : 한국이동통신 중앙연구소 사업지원연  
구팀 연구원

朴聖宇 (Seong Woo Park)

정회원

1962년 9월 13 일생

제19권 제2호 참조

金永川 (Young Chon Kim)

정회원

1956년 12월 10일생

현재 : 전북대학교 컴퓨터공학과  
부교수  
제19권 제2호 참조