

일반논문 (Regular Paper)

방송공학회논문지 제25권 제4호, 2020년 7월 (JBE Vol. 25, No. 4, July 2020)

<https://doi.org/10.5909/JBE.2020.25.4.609>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

영상 품질 및 전송효율 최적화를 위한 심층신경망 기반 영상전송기법

이 종 만^{a)}, 김 기 훈^{a)}, 박 현^{a)}, 최 증 원^{b)}, 김 경 우^{b)}, 배 성 호^{c)†}

Video Transmission Technique based on Deep Neural Networks for Optimizing Image Quality and Transmission Efficiency

Jong Man Lee^{a)}, Ki Hun Kim^{a)}, Hyun Park^{a)}, Jeung Won Choi^{b)}, Kyung Woo Kim^{b)},
and Sung Ho Bae^{c)†}

요 약

고품질 비디오 스트리밍 요구에 따라 제한된 대역폭에서 높은 전송률이 필요하고, 트래픽 혼재 상황이 더 발생한다. 특히 실시간 영상 서비스를 제공 시 패킷 손실 및 비트 오류 확률이 더 크게 증가한다. 이러한 문제를 해결하기 위해 실시간 서비스 품질향상을 위한 방법으로 FEC 기술의 한 종류인 랩터 코드가 어플리케이션 영역에서 활발히 사용되고 있다. 본 논문에서는 랩터 코드를 활용하여 유사한 수준의 화질에서 전송 효율을 높이기 위한 다양한 심층 신경망(Deep Neural Network, DNN) 기반 영상전송 파라미터를 결정하는 방법을 제안한다. 제안된 신경망은 패킷 손실율(Packet Loss Rate), 비디오 인코딩 속도 및 전송속도를 입력으로 사용하고 랩터 FEC 파라미터와 패킷 크기를 출력으로 한다. 제안한 방법은 기존 멀티미디어 전송 기법과 유사한 수준의 PSNR(Peak Signal-to-Noise Ratio)에서 전송 효율을 최적화하여 평균 1.2% 높은 스루풋(throughput)을 보였다.

Abstract

In accordance with a demand for high quality video streaming, it needs high data rate in limited bandwidth and more traffic congestion occurs. In particular, when providing real time video service, packet loss rate and bit error probability increase significantly. To solve these problems, a raptor code, which is one of FEC(Forward Error Correction) techniques, is pervasively used in the application layers as a method for improving real-time service quality. In this paper, we propose a method of determining image transmission parameters based on various deep neural networks to increase transmission efficiency at a similar level of image quality by using raptor codes. The proposed neural network uses the packet loss rate, video encoding rate and data rate as inputs, and outputs raptor FEC parameters and packet sizes. The results of the proposed method present that the throughput is 1.2% higher than that of the existing multimedia transmission technique by optimizing the transmission efficiency at a PSNR(Peak Signal-to-Noise Ratio) level similar to that of the existing technique.

Keyword : Real time video service, FEC(Forward Error Correction), DNN(Deep Neural Network), PSNR(Peak Signal-to-Noise Ratio)

1. 서론

네트워크 및 디지털 통신 기술의 발달로 음성 및 비디오와 같은 멀티미디어 데이터를 네트워크를 통해 실시간으로 전송할 수 있는 서비스에 대한 수요가 증가하고 있다. 현재 인터넷에서 비디오 콘텐츠 사용 비율은 전 세계 트래픽의 60%를 차지하며, 2020년에는 모든 인터넷 소비자 트래픽의 80%까지 점유할 것으로 예상된다^{[1][2]}.

이러한 요구사항을 충족하기 위해 CDMA2000 1x EV-DO, WiMAX, 3GPP LTE 및 5G와 같은 대용량 데이터 전송이 가능한 통신방식이 개발되었다^[3]. 그러나 트래픽 부하가 증가하고, 페이딩, 경로 감쇄, 신호 대 잡음 등 주변 환경에 따라 데이터 처리량이 급격히 저하 될 수 있기에 무선 네트워크 기술만으로 원활한 고품질 멀티미디어 서비스를 지원하는 것은 여전히 어려운 일이다^[4].

또한, 전송되는 데이터의 양을 줄이기 위해서는 효과적인 비디오 압축 알고리즘을 사용해야한다. 이를 위해 디지털 비디오 코딩 기술이 빠르게 발전하고 있으며, ISO/IEC 및 ITU에서 MPEG-2, MPEG-4, H.263, H.264, H.265 등과 같은 국제 표준을 확립하였다^[5].

특히, IP 계층 혼잡 및 비트 오류로 인해 패킷 손실은 실시간 비디오 스트림 품질에 굉장히 민감하다. 이러한 오류가 발생하기 쉬운 네트워크의 한계를 극복하기 위해 FEC 방식을 사용하여 서비스 품질을 개선하기 위한 노력이 이루어지고 있다^[6]. 일반적으로 비디오 전송 기술에 사용되는 응용계층에서의 FEC는 체계적인 특성을 가진 램터 코드이다. 램터 코드는 체계적인 코딩, 유연성, 코딩 효율성 및 속

도를 제공하는 블록 기반의 FEC 기술이며, 이를 기반으로 다양한 영상 스트리밍 기술들이 연구되어 왔다.

그러나 기존 멀티미디어 전송 기법의 큰 단점은 패킷 손실 및 전송속도, 비디오 인코딩에 따라 램터 FEC에서 다양한 패킷 크기를 실시간으로 변경할 수 없다는 것이다. 고정된 패킷 크기에서는 심볼 크기와 심볼 개수만 변경하여 인코딩 블록을 생성한다. 기존 램터 FEC 파라미터 결정방식은 패킷 손실율 및 전송속도, 그리고 비디오 인코딩에 따른 데이터 크기가 변하는 환경에서 이미지 품질과 처리량을 동시에 최적화하는 관점에서 한계가 있다. 사용 가능한 주파수의 제한성으로 인해 무선 자원의 효율적 사용 역시 고려할 요소이다.

본 논문에서는 패킷 손실, 전송 대역폭 및 비디오 인코딩 속도가 변경될 때 PSNR을 보장하는 범위 내에서 최대 스루풋(throughput)을 도출하는데 중점을 둔다. 또한, MPEG-4, H.264와 같은 고효율 압축 기술과 WiMAX 및 LTE와 같은 전송 트래픽을 증가시키는 무선 기술과 달리 패킷 크기와 램터 FEC 파라미터를 적응적으로 결정하여 전송효율을 최대화하는 DNN 구조를 제안한다. 딥 러닝 기반 RNN 또는 CNN 기술은 이미지 복원 또는 추적 분야에서 널리 사용되지만, 트래픽 전송 측면에서 응용 및 네트워크 계층 정보를 활용한 DNN 구조는 많지 않다. DNN 구조는 어플리케이션에 의해 제공되는 비디오 인코딩율, 네트워크 계층에 의해 제공되는 전송속도 및 패킷 손실율 기반으로 다중의 DNN 계층으로 구성된다. DNN 기반 영상전송기법은 다양한 비디오 인코딩율, 데이터 전송속도 및 패킷 손실을 적용하여 기존 멀티미디어 전송 기법과의 PSNR 및 스루풋(throughput) 성능을 비교 분석한다. 또한, 광범위한 테스트를 통해 멀티미디어 전송 효율의 극대화를 위한 최적의 DNN 구조를 도출한다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 램터 코드 구조 및 기존 멀티미디어 전송 기법과 2개의 입력 파라미터(패킷 손실율, 전송속도)로 구성된 기존 DNN 기반 영상전송기법을 검토하고, 3장에서는 DNN 기반 전송 파라미터 최적화 방법을 제안한다. 4장에서 최적의 DNN 구조를 식별하고 제안한 방법의 성능 검증을 수행하며 마지막으로 5장에서는 결론을 정리한다.

a) 한화시스템 (Hanwha Systems)

b) 국방과학연구소(Agency for Defense Development)

c) 경희대학교 컴퓨터공학과(Dept. of Computer Science and Engineering, Kyung-Hee University)

‡ Corresponding Author : 배성호(Sung-Ho Bae)

E-mail: shbae@khu.ac.kr

Tel: +82-031-201-2593

ORCID: <https://orcid.org/0000-0003-2677-3186>

※ Funding information : Agency for Defense Development (ADD),

Grant/Award Number: UC170024ED

• Manuscript received May 20, 2020; Revised July 7, 2020; Accepted July 9, 2020.

II. 관련 연구

1. 랩터 코드 소개

랩터 코드는 순방향 오류 정정 기술로서 체계화된(systematic) 코드이며, 유연성이 뛰어나고 높은 부호화 효율을 제공한다^[6]. 또한, 응용계층 순방향 오류 정정 코드로서 3GPP와 DVB-H에서 표준으로 채택되었다^[7]. 그림 1은 간단한 랩터코드의 구조이다. 랩터코드는 블록 단위로 인코딩이 이루어진다. 하나의 블록은 다시 여러 개의 소스 심볼들로 나누어지게 되고, 각 심볼들은 XOR 연산을 통해 인코딩 심볼을 생성하게 된다. 체계화된 랩터 코드의 인코딩 블록에는 소스 블록의 온전한 데이터와 오류 정정을 위한 여분의 데이터가 포함된다. 랩터 인코딩 블록의 소스 심볼 크기, 소스 심볼의 개수, 코드율은 예러 강인성, 인코딩/디코딩 소요 시간 및 스루풋(throughput)까지 영향을 미치게 된다.

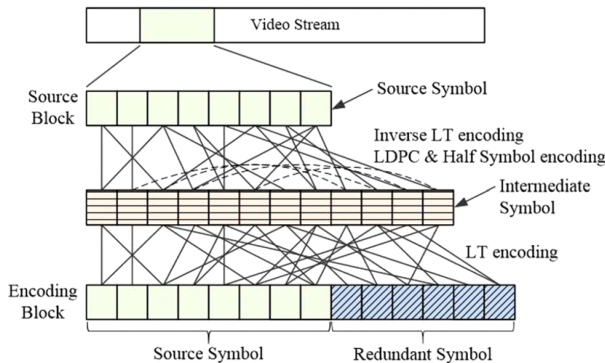


그림 1. 랩터코드 구조
Fig. 1. Raptor code structure

2. 기존 멀티미디어 전송 기법 소개

본 절에서는 기존 멀티미디어 전송 기법 구조에 대해 설명한다^[8].

비디오 인코더는 생성되는 데이터의 해상도, 프레임율 등의 코딩 파라미터를 FEC 파라미터 컨트롤 유닛에 제공하고, 또한 FEC 인코더 버퍼에 비디오 패킷을 전달한다. 랩터 인코더의 한 블록에 인코딩 가능한 비디오 데이터가 쌓이면 랩터 인코딩을 수행한다^[9]. FEC 파라미터 컨트롤 유닛에서는 네트워크 정보 즉, 패킷 손실율, 네트워크 지연

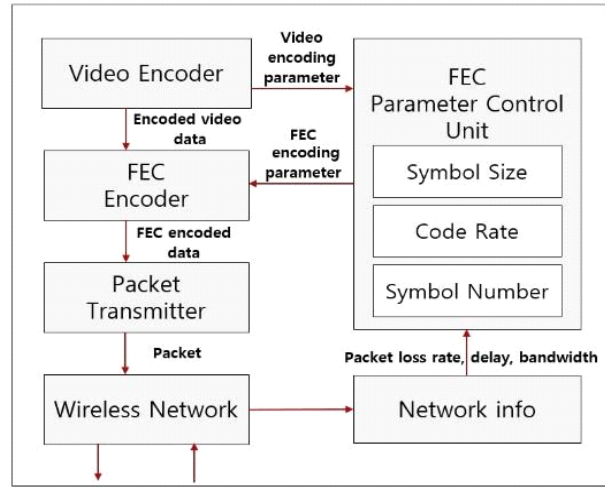


그림 2. 기존 멀티미디어 전송 기법 구조
Fig. 2. Existing multimedia transmission technique structure

및 가능한 전송속도를 기준으로 부호화율(c), 심볼 크기(s), 심볼 개수(k)를 실시간으로 결정한다^[9]. 심볼 크기와 심볼 개수는 랩터 인코딩/디코딩 복잡도에 큰 영향을 주기 때문에 네트워크 정보를 고려하여 결정되어야 하고, 코드율은 무선 채널 오류를 복원 할 수 있도록 결정되어야 한다. 패킷 크기는 심볼 크기와 심볼 개수의 조합으로 결정되며, 기존 멀티미디어 전송 기법에서는 고정하여 사용한다. 수신단에 도착한 패킷들은 역패킷화 과정과 랩터 디코딩 과정을 거쳐 재생 가능한 데이터로 변환된다. 이때 랩터 디코더는 체계화된 랩터 코드의 특성을 사용하므로 랩터 인코딩 블록내에 소스 심볼들이 온전하게 전송되었을 경우에는 디코딩 과정을 거치지 않는다.

3. 기존 멀티미디어 전송 기법 파라미터 결정 방법

본 절에서는 기존 멀티미디어 전송 기법에서 파라미터를 결정하기 위한 알고리즘을 간략하게 요약한다. 최적의 영상품질 및 전송효율을 위한 랩터 인코딩 블록의 심볼 크기(s), 심볼 개수(k), 그리고 코드율(c)를 결정한다.

정의한 문제의 해를 찾기 위해 전역 탐색 기반의 알고리즘을 사용한다. 즉, 알고리즘에서는 사용 가능한 (s,k)의 조합을 생성하고, 이 조합 내에서 스루풋(throughput)을 최대화하는 (s,k) 그리고 c를 탐색한다. 유한한 개수의 사용 가능한 (s,k) 조합을 제공하며, c는 주어진 (s,k)조합과 네트워크

상태 즉, 패킷 손실율(Packet Loss Rate, PLR)에 따라 결정 가능한 값이다. 파라미터 결정 알고리즘은 아래와 같다.

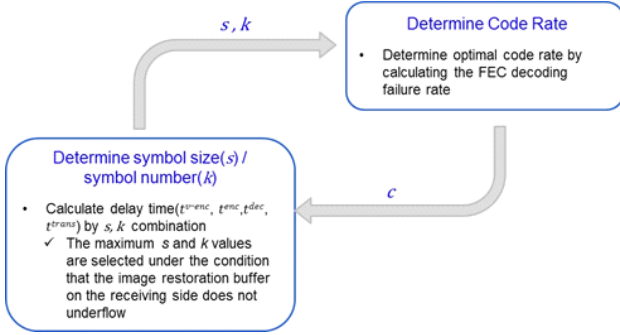


그림 3. AL-FEC 파라미터(s,k,c) 결정 방법

Fig 3. AL-FEC s, k, c parameter determination process

Step 1: 파라미터 후보가 저장되는 집합을 초기화 하고, 사용 가능한 모든 (s,k)의 조합 생성한다.

Step 2: 생성된 (s,k) 조합 중 하나를 선택하고, 선택된 (s,k) 기준으로 랩터 인코딩 블록에 해당되는 패킷의 개수 및 랩터 디코딩에 필요한 최소한의 패킷 개수, 패킷 손실율을 고려하여 코드율 c를 결정한다. 이때, $\Phi_i^{\max}$ 는 최대 허용 가능한 AL-FEC(Application Layer-Forward Error Correction) 디코딩 실패율이며, φ^{dec} 는 현재 AL-FEC 디코딩 실패율이다. AL-FEC 디코딩 실패율 계산은 이항 분포(Binomial Distribution)를 활용한다.

$$C^* = \arg \min C^{diff}(c) \quad 0 < c \leq 1 \quad (1)^{[8]}$$

$$C^{diff}(c) = \begin{cases} \Phi_i^{\max} - \varphi^{dec}(s, k, c) & \text{if } \varphi^{dec}(s, k, c) \leq \Phi_i^{\max} \\ \infty & \text{otherwise} \end{cases} \quad (2)^{[8]}$$

$$\begin{aligned} & \varphi^{dec}(s, k, c) \\ &= P(X < n^{\min}(s, k) \mid X \sim B(n^{enc}(s, k, c), 1 - plr)) \\ &= \sum_{i=0}^{n^{\min}(s, k)-1} \binom{n^{enc}(s, k, c)}{i} (1 - plr)^i (plr)^{n^{enc}(s, k, c)-i} \end{aligned} \quad (3)^{[8]}$$

Step 3: 선택된 (s,k) 조합에 대해 수신단 버퍼 언더플로우를 방지 할 수 있는지 계산한다.(수식 (4) 참고) t^{venc} 는 1개의 랩터 블록에서 비디오 인코딩이며, $t^{enc}(s, k, c)$ 와 $t^{dec}(s, k)$ 는 랩터 인코딩/디코딩 시간이다. $t^{trans}(s, k, c)$ 는 입력 데이터 속도와 관련된 랩터 인코딩 블록을 전송하는 시간이다. 마지막으로 buf^{recv} 는 수신 버퍼의 크기이다.

Step 4: step 3에서 결정된 (s,k) 조합과 c, 그리고 고정된 패킷 크기로 응용계층의 스루풋(throughput)을 계산한다. 즉, 심볼 크기와 심볼 개수의 조합으로 최대의 랩터 인코딩 블록크기를 찾는다.

$$\text{Throughput} = \frac{s \times k}{t^{venc} + t^{enc}(s, k, c) + t^{dec}(s, k) + t^{trans}(s, k, c)} \quad (5)$$

Step 5: 스루풋(throughput)을 최대로 하는 모든 (s,k) 조합에 대해 위 과정을 수행하였다면 파라미터 후보를 결정한다. 만약 아직 탐색하지 않은 (s,k) 조합이 있다면 step 2~4의 과정을 반복한다.

기존 멀티미디어 전송 기법의 가장 큰 단점은 랩터 인코딩 파라미터를 결정하는 알고리즘에서 패킷의 크기 변경이 불가하다. 패킷 크기는 심볼 크기와 심볼 개수 조합으로 고정되어 있다. 동일한 인코딩 블록 크기라고 가정하였을 때 패킷의 크기가 변경이 불가능하면, 패킷 개수도 고정되어야 한다. 앞서 설명하였듯이 패킷 개수는 인코딩 블록의 코드율 뿐만 아니라 스루풋(throughput)에도 큰 영향을 주는 파라미터이다. 1개의 인코딩 블록 기준으로 패킷 크기, 심볼 크기, 심볼 개수 조합은 반복연산 방식을 통해 결정되므로 시스템 연산에 부하가 클 것으로 판단된다. 이를 위해 본 논문에서는 심층 신경망 구조를 적용하여 인코딩 블록에서 학습된 데이터 기반으로 패킷 크기, 심볼 크기, 심볼 개수를 동적으로 변화하여 전송효율을 최대화한다.

$$t^{venc} + t^{enc}(s, k, c) + t^{dec}(s, k) + t^{trans}(s, k, c) \leq \frac{\chi \cdot buf^{recv}}{\text{Video decoding rate}}, \quad 0 \leq \chi \leq 1 \quad (4)$$

4. 기존 DNN 기반 영상전송기법

기존 DNN 기반 영상전송기법은 2개의 파라미터(패킷 손실율, 전송속도)를 DNN 구조의 입력으로 하고 DNN (32-32-16-16) 구조 통과 후, 3개의 파라미터(심볼 크기(s), 심볼 개수(k), 패킷 크기(p))를 출력한다^[10]. 기존 DNN 기반 영상전송기법은 출력 파라미터 대비 입력 파라미터 개수가 적어 DNN 성능 최적화의 한계가 있다. 또한, 응용계층에서 생성되는 비디오 인코딩율을 일정하게 고정하였다. 이를 보완하기 위해 본 논문에서는 DNN 입력 파라미터로 패킷 손실율 및 전송속도 뿐만 아니라 실시간으로 변화하는 비디오 인코딩율을 고려하여 최적의 DNN 구조 도출 및 전송 효율을 최대화 한다.

III. DNN 제안 구조

제안하는 DNN 구조는 아래 그림 4과 같다. 입력 파라미터는 비디오 인코딩율과 패킷 손실율, 전송속도이며, 출력 값은 랩터 인코딩 심볼 크기(s), 심볼 개수(k), 패킷 크기(p)이다. 제안 방법은 완전 연결 계층(fully connected layer)로 구성되며, 비선형성 확보를 위해 각각의 완전 연결 계층의 출력이 ReLU(Rectified Linear Unit)를 통과하도록 한다^[11]. ReLU는 음수 입력값을 0으로 맵핑하고 그렇지 않은 입력 값은 그대로 통과 시키는 맵핑 함수이다. ReLU를 통과한 i번째 완전 연결 계층은 아래와 같이 수식적으로 표현된다.

$$X_{i+1} = \max(W_i X_i + b_i, 0) \quad (6)$$

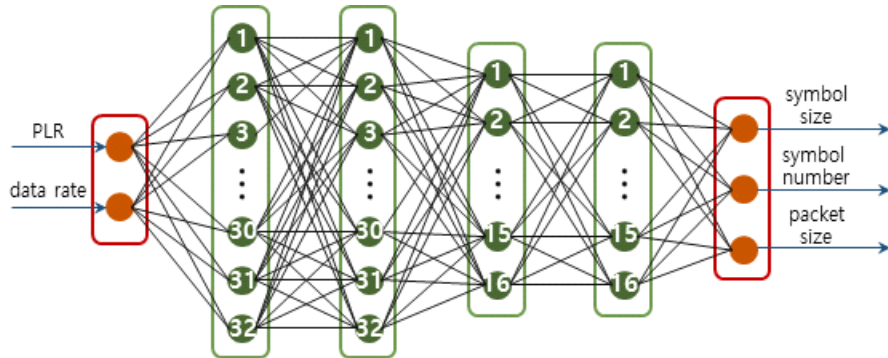


그림 4. 기존 DNN(32-32-16-16) 적용 구조
Fig. 4. Existing DNN application structure

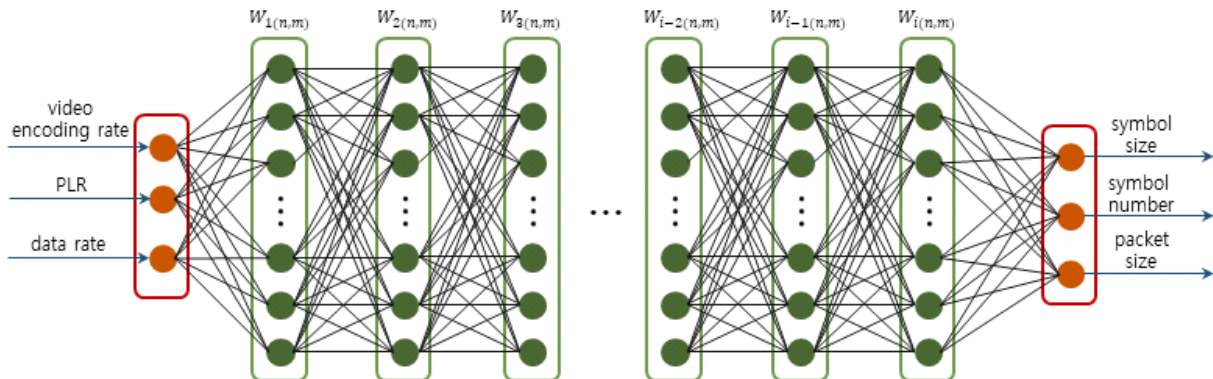


그림 5. DNN 제안 구조
Fig. 5. DNN proposal structure

여기서 $X_i \in R^{N \times 1}$ 및 $X_{i+1} \in R^{M \times 1}$ 은 i 번째 완전 연결 계층의 입력 및 출력, $W_i \in R^{M \times N}$ 및 $b_i \in R$ 는 완전 연결 계층의 모델 파라미터로 학습을 통해 결정된다. 우리는 광범위한 실험을 통해 최적의 심층신경망 파라미터/계층의 계수를 결정하였다.

IV. 시험결과

1. 시험데이터 획득

제안하는 DNN 구조를 검증하기 위해서는 기존 멀티미디어 전송 기법의 광범위한 결과가 필요하다. 전체 데이터 샘플을 추출하기 위해 안드로이드 임베디드 보드인 ODROID-U3를 사용하였다. 비디오 콘텐츠는 800x480 해상도, 25 프레임율을 적용하였다. 비디오 코덱은 H.264를 사용하며 1GOP(Group Of Picture)는 25개의 프레임(I-Frame : 1EA, P-Frame : 24EA)으로 구성된다. 네트워크 전송속도와 패킷 손실율은 네트워크 에뮬레이터를 활용하여 환경을 모사하였고, 비디오 인코딩율은 FFmpeg으로 변화를 주었다. AL-FEC 인코딩 블록은 32KB로 고정하고 수신단 버퍼 사이즈는 800ms 이하로 설정하였다.

앞서 설명한 것처럼 기존 멀티미디어 영상 기법에서는 패킷 크기가 고정되어 있다. 이에 각각에 패킷 크기(16/32/64/128byte) 기준으로 패킷 손실율, 전송속도, 비디오 인코딩율을 변화하여 최대의 스루풋(throughput)을 위한 (s,k,p)를 찾아낸다. 샘플 데이터 추출에 적용된 패킷 손실율, 전송속도, 비디오 인코딩율 범위는 표 1과 같다.

표 1. 비디오 인코딩율, 전송속도, 패킷 손실율 시험 범위/단위 간격
Table 1. Test range/unit steps in the video encoding rate, data rate, and packet loss rate

Database	Test Range	Unit Steps
video encoding rate	0.1~3.5Mbps	0.1Mbps
data rate	0.1~3.5Mbps	0.05Mbps
packet loss rate	0.1~5%	0.1% or 0.2%

2. 모델 파라미터 결정

제안한 DNN 구조의 계층 수와 모델 파라미터를 변경하는 광범위한 테스트를 통해 최적의 모델 파라미터를 찾는다. 테스트를 통해 추출한 5,700개의 샘플 중 학습셋으로 4,560개, 테스트셋으로 1,140개를 활용한다. 시뮬레이션을 위해 제안된 DNN 구조를 MATLAB 기반 딥 러닝 라이브러리에서 구현하였다. 이때, 비디오 인코딩율은 FFmpeg에서 실시간으로 변화하는 영상 인코딩율을 적용하였고, 전송속도 및 패킷 손실율은 MATLAB의 runif(n, min, max) 함수를 이용하여 균등 랜덤 분포(uniform random distribution)를 반영하였다. 제안 방법의 학습을 위해 ADAM 최적화 방법을 사용하였고^[12], 총 500번 세대(epoch)만큼 학습을 수행하였다. 학습률(learning rate)은 0.0001, 가중치 부식(weight decay)값은 0.0001을 사용하였다. 학습 시 미니 배치의 크기는 32로 설정하였다. 학습을 위한 비용 함수는 예측된 $(\hat{s}, \hat{k}, \hat{p})$ 와 실제 값 s, k, p 간 평균 자승 오차(Mean Squared Error; MSE)를 사용하였다. (7)는 비용함수(J)의 수식적 표현이다.

$$J = \frac{1}{n} \sum_{i=0}^n (s_i - \hat{s}_i)^2 + (k_i - \hat{k}_i)^2 + (p_i - \hat{p}_i)^2 \quad (7)$$

여기서 s, k, p가 가질 수 있는 값은 각각 $s=\{16, 32, 64, 128\}$, $k=\{128, 256, 512\}$, $p=\{16, 32, 64, 128\}$ 이다. 각 파라미터 값의 동적 범위가 다르면, 결과적으로 비용함수를 최적화 시 각 파라미터 마다 다른 가중값을 가지게 된다. 이를 해결하기 위해 로그(log) 스케일로 치환된 s, k, p 값에 대해 학습을 수행하였다.

(2, 4, 8)개의 계층 수와 (128, 64, 32, 16, 8, 4)의 각 계층에 대한 뉴런의 수에 대해 실험을 수행하였다. 일반적인 DNN 구조는 하위 계층에서 상위 계층으로 갈수록 차원이 축소(dimensionality reduction)된다^{[13][14][15]}. 본 논문도 마찬가지로 계층이 상위로 갈수록 뉴런의 수가 감소하는 형태로 모델을 구성하였다. 표 2는 다양한 모델 파라미터에 대한 테스트 데이터셋에서의 MSE 성능을 제시한다. 본 논문에서는 표 2의 결과를 기반으로 MSE 최소화 관점에서 (16-16-8-8) DNN 구조를 최종 제안 모델로 선택하였다.

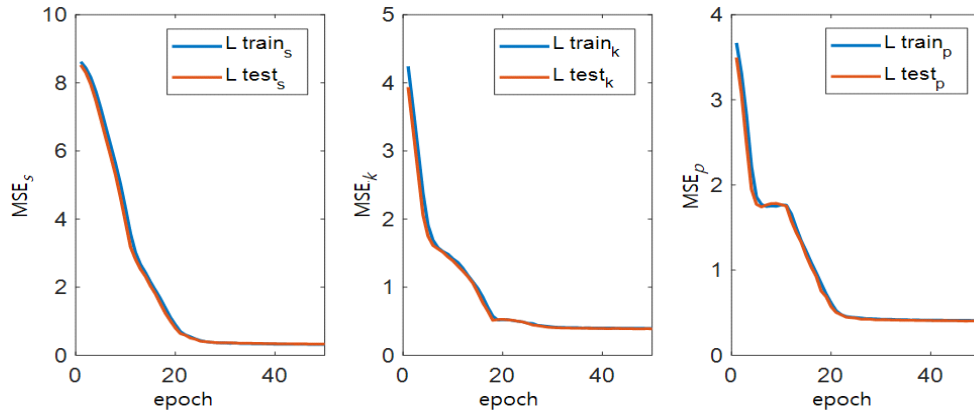


그림 6. 심볼 크기/심볼 개수/패킷 크기의 Training 및 Testing MSE

Fig. 6. Training and testing performance on mean squared error (MSE) for symbol size (s), symbol number (k), and packet size (p) along the number of epochs

표 2. 다양한 계층 및 뉴론 수에 따른 testing dataset의 MSE 성능

Table 2. MSE performance on the testing dataset under various model parameter conditions.

Number of neurons per layer	MSE			
	S	k	P	S+k+P
8-4	0.2937	0.3775	0.3926	1.0638
16-8	0.2957	0.3734	0.3901	1.0592
32-16	0.2956	0.3727	0.3899	1.0582
64-32	0.3053	0.3781	0.3954	1.0788
128-64	0.2994	0.3752	0.3926	1.0671
8-8-4-4	0.2941	0.3727	0.3891	1.0559
16-16-8-8	0.2911	0.3719	0.3885	1.0515
32-32-16-16	0.3062	0.3774	0.3933	1.0769
64-64-32-32	0.3026	0.382	0.3968	1.0814
128-128-64-64	0.3162	0.3934	0.3958	1.1054
8-8-8-8-4-4-4-4	0.2946	0.3794	0.3916	1.0657
16-16-16-16-8-8-8-8	0.315	0.3796	0.3993	1.0939
32-32-32-32-16-16-16-16	0.3003	0.3771	0.3932	1.0706
64-64-64-64-32-32-32-32	0.3029	0.3781	0.3943	1.0753
128-128-128-128-64-64-64-64	0.3042	0.3808	0.3955	1.0804

제안한 (32-32-16-16) DNN 구조의 성능을 평가하기 위해 k-폴드 교차 검증(k-fold cross validation)을 수행하였다. k = 5로 설정하여 총 5700개의 데이터에서 80%는 Training set, 20%는 Testing set으로 활용하였다. 교차 검증 시에는 Testing set이 중복되지 않도록 5개의 set으로 분류하여 테스트를 수행하였다. 각 Testing set에서 예측한 심볼 크기, 심볼 개수, 패킷 크기를 기반으로 스루풋(throughput)을 산정한다. 그리고 기존 멀티미디어 전송 기법의 스루풋(through-

put)은 k-폴드 교차 검증 수행 시 분류된 Testing set의 패킷 손실율과 전송속도, 비디오 인코딩을 기준으로 패킷 크기 (16/32/64/128byte)를 고정하여 각각 산정하였다.

결론적으로 기존 멀티미디어 전송 기법에서는 패킷 크기를 고정하여 파라미터를 결정하고, 심층 신경망 구조에서는 학습기반 데이터로 패킷 크기 및 램퍼 FEC 파라미터(심볼 크기, 심볼 개수)를 적응적으로 결정한다.

3. 결과분석

제안한 DNN 구조의 스루풋(throughput)은 최종 선정된 DNN 구조(16-16-8-8)를 사용하여 예측된 s , k , p 값을 사용하여 산출된다. 기존 멀티미디어 전송 기법의 스루풋(throughput)은 고정 패킷 크기 16/32/64/128byte에서 s 및 k 값을 기반으로 계산된다. 스루풋(throughput)을 산출하는 수식은 기존 멀티미디어 전송 기법의 공식과 동일하다.

기존 멀티미디어 전송 기법은 리눅스 기반의 C언어로 동작하여 파라미터를 도출하고, 제안한 심층 신경망은 MATLAB을 통해 파라미터를 결정한다. 동일한 환경에서 성능 비교/분석을 수행하지는 않았지만, 하나의 인코딩 블록 기준으로 두 환경의 파라미터 결정 알고리즘 수행시간 차이가 30ms 이하로 확인된다. 이는 실제 스루풋(throughput) 성능에는 영향이 미비하므로 두 알고리즘의 수행시간의 차이는 스루풋(throughput) 도출 시 고려하지 않았다. 본 논문의 MATLAB 시뮬레이션에서 적용한 오류 확률은 균등 랜덤 분포(uniform random distribution)이며, 송신측에

서는 CSI(Channel State Information)를 기반으로 획득된 패킷 손실율을 다음 랩터 인코딩 블록 생성 시 반영하는 것으로 가정하였다. 단, 실제 환경에서는 적용되는 모델 특성에 따라 송신측으로 전달되는 패킷 손실율 주기 및 랩터 인코딩 적용 시점이 달라 스루풋(throughput)의 차이는 발생할 여지가 있다. 표 3은 제안한 (16-16-8-8) DNN과 기존 멀티미디어 전송 기법의 성능 비교 결과이다.

제안된 DNN 구조를 적용하면 기존의 멀티미디어 전송 기법에 비해 스루풋(throughput)이 향상되었다. 제안된 DNN 아키텍처는 스루풋(throughput)을 최대화하기 위해 패킷 크기와 랩터 FEC 파라미터를 모두 최적화하였다. 기존의 멀티미디어 전송 기법에서 랩터 FEC 파라미터는 고정된 패킷 크기를 갖는 패킷 손실율, 전송속도 및 비디오 인코딩율에 따라 결정되는 반면, 제안된 DNN 구조는 학습된 데이터 베이스 기반으로 하여 패킷 크기 및 FEC 파라미터를 적응적으로 결정한다. 기존 멀티미디어 전송 기법 대비 패킷 크기를 가변함으로써 1개의 랩터 인코딩 블록 생성 시 송신 가능한 전송속도에 최대한으로 근접한 랩터 인코

표 3. 기존 멀티미디어 전송 기법과 제안 DNN구조의 성능 비교 결과

Table 3. Performance of existing multimedia transport technique and the proposed DNN structure

구분	Testing Dataset	Packet Size (byte)	Throughput (kbps)	Avg. PSNR
Proposed DNN structure	k=1	Adaptive (16/32/64/128)	207.88	30.55
	k=2		207.40	
	k=3		205.15	
	k=4		207.67	
	k=5		208.79	
Existing multimedia transport technique	k=1	16	204.73	31.22
	k=2		203.99	
	k=3		201.96	
	k=4		204.67	
	k=5		208.79	
	k=1	32	206.29	31.23
	k=2		205.77	
	k=3		203.60	
	k=4		206.69	
	k=5		207.35	
	k=1	64	206.04	31.32
	k=2		205.56	
	k=3		203.38	
	k=4		205.85	
	k=5		206.91	
	k=1	128	204.58	31.13
	k=2		203.63	
	k=3		201.51	
	k=4		203.94	
	k=5		205.01	

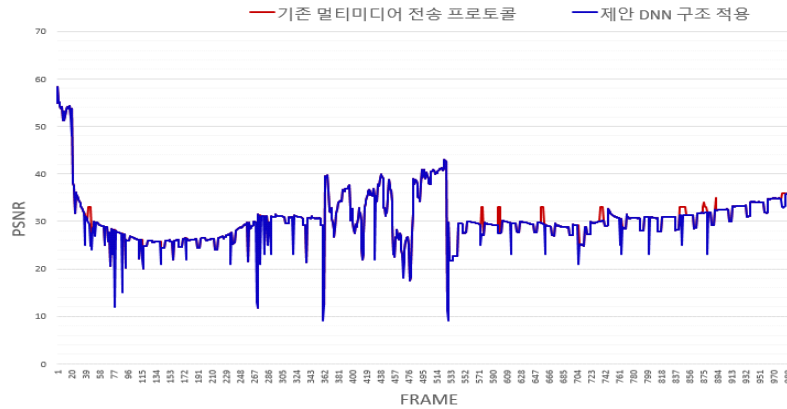


그림 7. 기존 멀티미디어 전송 기법과 제안 DNN구조의 PSNR 성능 비교
Fig. 7. PSNR performance comparison between the existing multimedia transport technique and proposed DNN structure

딩 블록 크기로 생성이 가능하기 때문이다. DNN은 다양한 비선형 함수를 근사화 할 수 있는 특성(Universal Approximation Theorem)을 가지고 있기 때문에 산술적 반복 연산을 적용한 기존 멀티미디어 전송 기법 대비 최적의 해를 도출하는데 효율적이다^[16]. 즉, 최대의 스루풋(throughput)을 위한 조합(패킷 크기, 심볼 크기, 심볼 개수)을 적용할 수 있는 구조이다. 따라서, 랜덤하게 변하는 채널 환경에 신경망 구조를 적용함으로써 수행 될 수 있는 적절한 패킷 크기를 선택함으로써 최대 스루풋(throughput)이 개선 될 수 있다.

또한, 제안된 신경망 구조에서의 PSNR은 기존의 멀티미디어 전송 기법과 비교하여 약간의 성능 차이를 보여준다. DNN 구조에서 PSNR이 떨어지는 경우는 잘못된 조합을 선택하여 랩터 디코딩이 실패한 경우로 보여진다. 이는 패킷 크기 및 랩터 인코딩 파라미터 예측을 위한 충분한 입력 정보가 제공되지 않은 DNN 모델의 한계라고 판단된다.

V. 결 론

본 논문에서는 비디오 품질 및 비디오 전송 효율을 최적화하기 위한 16-16-8-8 DNN 기반 전송 파라미터 추정 방법을 제안하였다. 기존의 멀티미디어 전송 기법에서는 패킷 크기가 고정된 상태에서 랩터 FEC의 파라미터가 결정

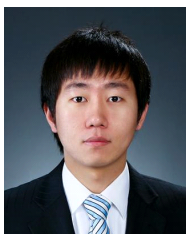
되었다. 기존 멀티미디어 전송 기법의 단점을 극복하기 위해 비디오 인코딩을, 데이터 전송속도 및 패킷 손실율에 따라 적절한 패킷 크기를 선택하여 스루풋(throughput)을 개선한 DNN 구조를 도출하였다. 광범위한 테스트 데이터 분석 및 시뮬레이션을 통해 DNN 구조는 기존 멀티미디어 전송 기법 대비 스루풋(throughput)이 1.2% 향상됨을 확인하였다. 다시 말해서 제안된 방법은 동일한 네트워크 대역폭에서 유사한 레벨의 PSNR로 더 많은 비디오 데이터를 전송할 수 있다. 하지만, 이미 기존 멀티미디어 전송 방법이 최적에 가까운 성능을 도출하기에 DNN이 향상시킬 수 있는 성능이 많지 않다고 판단된다. 그럼에도 불구하고 스루풋(throughput)이 향상된 것은 universal approximation theorem에 따라 non-linear 함수를 더욱 잘 모델링했기 때문이라고 할 수 있다. 또한, PSNR이 떨어지는 경우는 DNN에서 잘못된 조합을 선택하여 랩터 디코딩이 실패한 경우로 PSNR이 다소 떨어진 것으로 보여진다. 이는 패킷 크기 및 랩터 인코딩 파라미터 예측을 위한 충분한 입력정보가 제공되지 않은 DNN 모델의 한계라고 판단된다. 추후, DNN 입력 파라미터로 SNR(Signal to Noise Ratio, RSSI(Received Signal Strength Indicator) 등 무선채널환경을 보다 정확하게 반영할 수 있는 값을 활용하면 잘못된 조합을 개선할 수 있다고 본다. 주파수 자원은 제한되어 있고 최근 트래픽 양은 증가하는 상황에서 대역폭을 효율적으로 사용하는 것이 중요하다. 따라서 제안된 방법을 적용하면 네트워크 리

소스가 제한되거나 트래픽 환경이 정체된 경우 비디오 전송을 보다 효율적으로 향상시킬 수 있다. 향후 연구에서는 입력 파라미터를 추가한 제안 DNN 알고리즘 및 웨이브폼을 실장비에 탑재하고, 네트워크 채널 변경에 따른 패킷 손실을 반영하여 TTF(Time to Failure)를 분석하도록 하겠다.

참 고 문 헌 (References)

- [1] Cisco. White Paper, "Cisco VNI Forecast and Methodology," 2015 - 2020.
- [2] SANTOS-GONZÁLEZ, Iván, et al. Implementation and analysis of real-time streaming protocols. *Sensors*, 2017, 17.4: 846.
- [3] Kwon, Oh Chan, Yunmin Go, and Hwangjun Song. "An Energy-Efficient Multimedia Streaming Transport Protocol Over Heterogeneous Wireless Networks." *IEEE Trans. Vehicular Technology* 65.8 (2016): 6518-6531.
- [4] Motorola White Paper, "Realistic LTE Performance from Peak Rate to Subscriber Experience," 2009.
- [5] M. Luby, "LT codes," in *Proceedings of the Annual Symposium on Foundations of Computer Science*, pp. 271-280, 2002.
- [6] A. Shokrollahi, "Raptor codes," *IEEE Tr. on Information Theory*, Vol. 52, No. 6, pp. 2551 - 2567, June, 2006.
- [7] O. C. Kwon and H. J. Song, Raptor Codes-based Multipath Multimedia Transport Protocol over Wireless Networks, *Proceedings of the Korea Information Science Society Conference*, 1243-1245, 2014
- [8] GO, Yunmin; SONG, Hwangjun. Raptor codes-based energy-efficient screen mirroring system. In: *Visual Communications and Image Processing (VCIP)*, 2016. IEEE, 2016. p. 1-4.
- [9] KWON, Oh Chan, et al. MPMTTP: Multipath multimedia transport protocol using systematic Raptor codes over wireless networks. *IEEE Transactions on Mobile Computing*, 2015, 14.9: 1903-1916.
- [10] Jong-Man Lee, (2018). A New Parameter Determination Method for Image Transmission protocol Using Deep Neural Network. *Journal of the Institute of Electronics Engineers*, 55(1), 30-36.
- [11] Nair, Vinod, and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines." *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010.
- [12] Kingma, Diederik, and Jimmy Ba. "Adam: A method for stochastic optimization." , Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015
- [13] Y. Bengio, A. Courville, and P. Vincent., "Representation Learning: A Review and New Perspectives," *IEEE Trans. PAMI*, special issue Learning Deep Architectures, 2013
- [14] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview", *Neural Networks*, Vol 61, pp 85-117, Jan 2015
- [15] I. J. Kim, Deep Learning: A new trend in machine learning. *Journal of the Korean Institute of Communication Sciences (Information and Communication)*, 31 (11), 52-57, 2014
- [16] Csáji, Balázs Csanád. "Approximation with artificial neural networks." *Faculty of Sciences, Eötvös Loránd University, Hungary* 24.48 (2001)

저 자 소 개



이 종 만

- 2010년 2월 : 경희대학교 전자전파공학과 학사
- 2010년 1월 ~ 현재 : 한화시스템 전문연구원
- 2020년 3월 ~ 현재 : 연세대학교 전기전자공학과 석사과정
- ORCID : <http://orcid.org/0000-0002-8467-6504>
- 주관심분야 : 전술통신, 이동통신, 모바일 단말, 딥러닝



김 기 훈

- 2002년 2월 : 명지대학교 정보통신공학 학사
- 2015년 2월 : 아주대학교 IT융합공학 석사
- 2008년 1월 ~ 현재 : 한화시스템 수석연구원
- ORCID : <http://orcid.org/0000-0002-3519-0268>
- 주관심분야 : 무선통신시스템, 국방전술 네트워크, IoT, SDN

저 자 소 개



박 현

- 2005년 2월 : 광운대학교 컴퓨터공학부 학사
- 2007년 2월 : 광운대학교 컴퓨터공학과 석사
- 2010년 4월 ~ 현재 : 한화시스템 전문연구원
- ORCID : <http://orcid.org/0000-0003-2643-3130>
- 주관심분야 : 군 전송통신망, 모바일 네트워크, SDN



최 증 원

- 1989년 2월 : 충남대학교 계산통계학과 학사
- 1993년 8월 : 충남대학교 계산통계학과(전산학) 석사
- 1997년 8월 : 충남대학교 전산학과 박사
- 1997년 7월 ~ 현재 : 국방과학연구소 수석연구원
- 2013년 9월 ~ 현재 : 과학기술연합대학원대학교 부교수
- ORCID : <http://orcid.org/0000-0002-3642-2323>
- 주관심분야 : 전송통신, 위성통신, 인지무선통신, 바이오통신, 정보융합 등



김 경 우

- 2000년 2월 : 경북대학교 전자공학과 학사
- 2002년 2월 : 경북대학교 전기전자공학부 석사
- 2005년 4월 : MS, Dept. ECE, University of Florida
- 2012년 5월 : PhD, Dept. ECE, University of Florida
- 2013년 9월 ~ 현재 : 국방과학연구소 선임연구원
- ORCID : <https://orcid.org/0000-0003-1797-1869>
- 주관심분야 : 전송통신, 인지무선통신, 확률기반 데이터 분석 등



배 성 호

- 2011년 : 경희대학교 학사
- 2012년 : KAIST 전기 및 전자공학 석사
- 2016년 : KAIST 전기 및 전자공학 박사
- 2016년 ~ 2017년 : MIT 컴퓨터과학 박사 후 과정
- 2017년 ~ 현재 : 경희대학교 컴퓨터공학과 조교수
- ORCID : <http://orcid.org/0000-0003-2677-3186>
- 주관심분야 : 심층신경망 모델 압축/해사/탐색, 이미지 신호의 역문제