한국항행학회 논문지 제13권 제4호 2009년 8월

적응형 화면 스킵 기반 실시간 비디오의 소스/채널 통합 부호화율 제어

Joint Source/Channel Rate Control based on Adaptive Frame Skip for Real-Time Video Transmission

이명진*

Myeong-Jin Lee*

요 약

본 논문에서는 패킷 손실 채널로의 전송을 위한 비디오 부호기의 소스/채널 부호화율 통합 제어 방식을 제안한다. 비디오 전송 시스템의 지연 제약에 따른 소스 부호화율 상한을 구하고, 부호화되거나 스킵된 화면에 대한 소스 왜곡과 패킷 손실에 의한 채널 왜곡을 기반으로 일정 부호화 구간에 대한 평균 왜곡을 정의한다. 주어진 채널률과 패킷 손실률에 대해 비디오 부호화 계수와 FEC 블록 크기를 제어계수로 하는 평균 왜곡 최소화 문제를 정의하고, 실시간 해결 알고리즘을 제안한다. 모의실험에서 제안방식은 40% 이하의 패킷 손실률과 30 프레임 이상의 인트라 매크로블럭 갱신률에서 TMN8에 비해 PSNR과 주관적 화질에서 우수한 성능을 보였다.

Abstract

In this study, we propose a joint source/channel rate control algorithm for video encoder targeting packet erasure channel. Based on the buffer constraints of video communication systems, encoding rate constraint is presented. After defining source distortion models for coded and skipped video frames and a channel distortion model for packet errors and their propagation, an average distortion model of received video is proposed for a given encoding window. Finally, we define an optimization problem to minimize the average distortion for given channel rates and packet loss rates by controlling spatio-temporal parameters of source video and FEC block sizes. Then, we propose a window-based algorithm to solve the problem in real-time.

Key words : joint source/channel rate control, FEC, video rate control, video transmission system

I. 서 론

최근들어 인터넷과 무선망을 통한 비디오 서비스에 대한 수요가 급증하고 있어서 손실 채널을 통한 적응형 비디오 전송 기법에 대한 연구가 활발히 진행되고 있다. 손실 채널상에는 비트 에러나 패킷 손실을 피할 수 없기 때문에, 비디오 부호기는 에러에 강인한 부호화 기법, 에러 은닉 및 복원 기법 등을 활용하여 이에 대처하고 있다. 그러나, 이들 기법들은 지연, 대역폭 등에 대한 오버헤드가 존재하고 이는 수

* 한국항공대학교 항공전자및정보통신공학부(School of Electronics, Telecommunications and Computer Engineering, Korea Aerospace University)
· 제1저자 (First Author) : 이명진
· 투고일자 : 2009년 6월 29일
· 심사(수정)일자 : 2009년 6월 30일 (수정일자 : 2009년 7월 22일)
· 게재일자 : 2009년 8월 30일

신 비디오의 화질 저하로 이어질 수 있다.

저지연의 실시간 비디오 전송을 위해서는 전송단과 수신단 간의 패킷 손실 정보를 교환하여 손실 패킷을 복원하는 ARQ (Automatic repeat request) 기법보다는 채널 상황에 맞게 FEC (Forward Error Correction)를 적용하여 전송하는 방식이 일반적이다. FEC 기법은 소스 패킷에 중복성을 더하여 수신단에서 일부 패킷의 손실이 존재해도 원래의 패킷 그룹을 복원할 수 있으나, 패리티 패킷 생성시 채널 비트 자원을 사용하므로 소스 부호기에서 사용 가능한 채널 비트 자원은 감소한다. 즉, 제한된 가용 채널 대역폭 하에서 패킷 손실에 의한 비디오의 왜곡 증가를 FEC에 의해 감소시킬 수 있으나, 소스 비디오 왜곡은 증가하게 된다. 수신단의 왜곡은 소스 비디오 왜곡과 채널 왜곡의 중첩되어 계산될 수 있고, 제한된 채널 자원하에서 이들 왜곡의 합을 최소화할 수 있는 비디오소스 부호화율 제어와 패킷화 및 채널 부호화 제어가 필요하다.

비디오 부호기에서의 부호화율 제어는 부호기의 출력 비트열의 채널 적응 여부와 수신 화질에 큰 영향을 미칠 수 있다. 기존의 연구들은 대부분 고정된 비디오 화면률과 공간 해상도에 대한 부호화율 제어 방식에 집중되어 왔다. 그러나, 손실 채널 환경에서는 비디오 소스의 부호화 복잡도와 소스 부호기에서 사용 가능한 채널 대역폭이 시변하기 때문에 비디오의 총 왜곡은 단순히 양자화 계수나 화면률 같은 단일의 부호화 계수로는 최소화가 불가능하다. 최근에는 비디오의 시간방향 및 공간방향 화질을 동시에 고려한 부호화율 제어 알고리즘들이 제안되었다[1, 2, 3]. 그러나, 비디오 소스의 사전 분석을 위해 추가의 버퍼 지연이 필요하거나 최적의 해를 얻기 위해 동적 프로그래밍 기법이 사용되어서 영상전화나 영상회의와 같은 실시간 비디오 서비스에는 적용이 불가능하다[2, 3]. [1]에서는 부호화된 화면과 스킵된 화면들에 대한 명시적인 왜곡 함수를 이용하여 시공간 최적화 문제를 단순화시켰으나 스킵된 화면에 대한 왜곡 모델이 정확하지 않고, 종단간 지연 문제를 부호화율 제어시 고려하지 않았다. [4]에서는 현재까지 입력된 비디오 소스의 특성과 부호화된 화면들의 왜곡 특성을 기반으로 스킵되거나 부호화될 화면의 왜곡 모델을 제시하였고, 종단간 지연을 고려한 부호화율 제어 방식을 제안하였다. 그러나, 이 방식 역시 손실 채널상의 비트 오류나 패킷 손실에 대한 고려는 이루어지지 않았다.

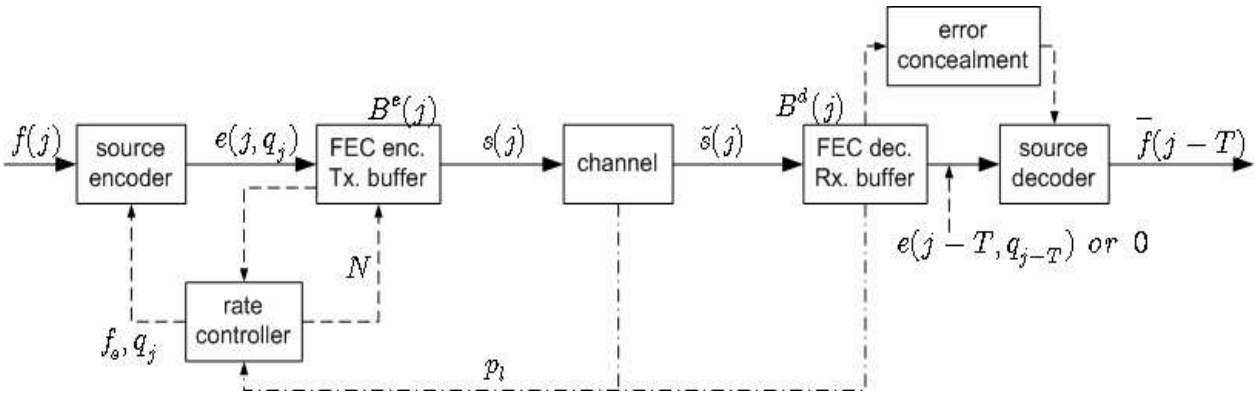

그림 1. 소스/채널 통합 부호화 기반 비디오 전송 시스템

Fig. 1. Video transmission system of joint source/channel rate control

본 연구에서는 손실 채널에서 비디오의 수신 왜곡을 최소화하기 위해 소스/채널 통합 부호화율 제어 알고리즘을 제안한다. 제 2장에서는 종단간 지연 제약을 기반으로 부호화율 상한을 유도하고, 비디오 패킷화 방식과 패킷 기반 채널 부호화 방식을 설명한다. 제 3장에서는 비디오 소스 부호기에서 스킵된 화면과 부호화된 화면의 왜곡 모델을 제시하고, 채널 에러 전파에 따른 수신단의 채널 왜곡 모델을 설명한다. 제 4장에서는 적응형 화면 스킵 기반 소스/채널 통합 부호화 문제를 정의하고, 실시간 알고리즘을 제안한다. 제 5장에서는 제안된 방식의 성능 평가를 위한 모의실험 환경을 설명하고, 실험 결과를 제시한다. 마지막으로 제 6장에서 결론을 맺는다.

Ⅱ. 비디오 전송 시스템

본 절에서는 그림 1의 손실 채널에서 실시간 비디오 전송 시스템의 구조를 제시하고, 종단간 지연 제약 기반의 버퍼 제약 조건을 유도한다. 또한, 비디오의 패킷화 과정과 이를 기반으로 한 채널 부호화 방식을 설명하고, 패킷 손실에 따른 수신 비디오의 성능 분석을 위한 도구를 제시한다.

2-1. 비디오 전송 시스템의 버퍼 제약

비디오 부호기의 버퍼 레벨은 다음과 같다.

$$B^e(j)=\max\{B^e(j-1)+e(j,q_j)-s(j),0\} \quad (1)$$

여기에서 $B^e(j)$, $e(j,q_j)$, $s(j)$는 j번째 화면의 부호기 버퍼 레벨, 양자화계수 q_j로 부호시 발생 비트량, 부호기 버퍼의 서비스율이다. 부호기 버퍼 크기는 충분히 크다고 가정한다.

실제 부호기 버퍼의 서비스율은 다음과 같다.

$$\tilde{s}(j)=\min\{s(j),B^e(j-1)+e(j,q_j)\}. \quad (2)$$

복호기 버퍼 레벨은 다음과 같이 표현될 수 있다.

$$B^d(j)=\begin{cases}\sum_{i=1}^{j}\tilde{s}(i)-\sum_{i=1}^{j-T}e(i,q_i), & j\geq T\\ \sum_{i=1}^{j}\tilde{s}(i), & j<T\end{cases} \quad (3)$$

여기에서 T는 부호기 버퍼 입력 시점부터 복호 시점까지의 종단간 지연을 나타낸다.

부호기 버퍼 레벨은 식 1, 2를 이용하면 $B^e(j)=\sum_{i=1}^{j}\{e(i,q_i)-\tilde{s}(i)\}$ 로 표현될 수 있다. 식 3에 복호기 버퍼 고갈 조건($B^d(j)\geq 0$)을 적용하면, 복호기 버퍼 고갈 조건은 $B^e(j)\leq\sum_{i=j+1}^{j+T}\tilde{s}(i)$ 와 같이 유도된다. 부호기 버퍼에 버퍼링된 비디오 데이터는 최대 가용 채널률 $s(j)$로 전송되기 때문에, 식 1, 3과 결합한 복호기 버퍼 고갈 조건은 다음과 같이 부호기 발생 비트율의 조건으로 표현될 수 있다.

$$e(j,q_j)\leq\sum_{i=j}^{j+T}s(i)-B^e(j-1). \quad (4)$$

2-2. 손실 채널 상의 비디오 패킷화 및 채널 부호화 방식

네트워크 상의 패킷 손실을 복원하기 위한 FEC 기법은 오류 정정을 위한 지연이 ARQ보다 적어서 실시간성이 요구되는 비디오 응용에 적합하다. Reed-Solomon 부호와 같은 FEC 기법이 주로 사용되는데, 전송단은 매 K개 소스 패킷들에 대해 $(N-K)$개 패리티 패킷들을 생성하여 N개 패킷들로 구성된 FEC 블록을 구성한다. 수신단에서는 N개 패킷들 중에서 최소한 K개 이상의 패킷들이 수신되면 모든 K개의 소스 패킷들을 복원해 낼 수 있다.

손실 채널 상에서의 소스 비디오 패킷들의 FEC 부호화시 비디오 부호화 및 패킷화 구조가 고려되어야 한다. H.263+ 비디오 압축 비트열의 경우 패킷은 한 개 이상의 GOB(Group of Blocks)들로 구성된다. 본 연구에서는 한 개의 GOB를 하나로 패킷화하였고, RS(N, K) 채널 부호기를 이용하여 매 K개 패킷마다 $(N-K)$개 패리티 패킷을 생성한다. QCIF 비디오에 대해 채널 부호화는 K를 9로 설정하여 화면 단위로 수행하였다. 본 연구에서는 화면 내의 GOB단위 패킷들의 평균 크기를 기준으로 FEC 블록을 구성하였다.

전송 채널은 패킷 손실 확률 p_l의 베르누이(Bernoulli) 패킷 손실 모델로 가정되었다. 이 경우 채널 복호후 잉여 에러 확률(residual error probability)는 다음과 같이 기술될 수 있다.

$$P_L(N)=\sum_{k=N-K+1}^{N}\binom{N}{k}p_l^k(1-p_l)^{N-k} \quad (5)$$

식 5는 FEC 블록 전체의 손실 확률을 나타낸다. 따라서, 비디오 화면 단위로 FEC 블록을 구성하는 경우 식 5는 화면 전체의 손실 확률을 나타낸다. 그러나, [5]와 같이 FEC 블록의 복원이 불가능하더라도 오류가 없는 소스 패킷을 복호에 활용할 수 있는 경우에는 다음과 같은 소스 패킷 손실 확률을 구하여

평균 PSNR을 구해야 할 것이다.

$$p_l^*(N) = p_l P_L(N-1) \tag{6}$$

본 연구에서는 압축 비디오 비트열의 패킷화 방법은 일정하다고 가정하고, 주어진 채널 제약하에서 최적의 화질을 얻을 수 있는 비디오의 시공간 방향 부호화 계수와 FEC 블록 크기 제어를 목표로 한다. 따라서, 보다 분석이 용이한 화면 단위의 FEC 부호화를 수행하고 식 5를 이용하여 화면의 손실 또는 복호화 여부를 판단하였다.

Ⅲ. 소스/채널 부호화 계수에 따른 비디오 왜곡 모델

기존의 소스/채널 통합 부호화율 제어에서는 단순히 소스 부호기의 부호화율을 양자화 계수를 이용하여 조절하였기 때문에 채널 에러율이 높은 경우에는 급격한 화질 저하가 빈번히 발생하였다. 본 연구에서는 이와 같은 문제를 해결하기 위해 채널 상황에 적응적으로 비디오 화면을 스킵하거나 양자화 계수를 제어하는 방법을 사용한다. 수신 비디오의 왜곡을 계산하여 화면 스킵 여부나 양자화 계수를 결정하기 위해서는 시공간방향 부호화 계수에 따른 비디오 왜곡 모델이 필요하다. 이를 위해 이전 연구 결과를 본 연구에 적용한다[4].

3-1. 참조 화면과의 거리에 따른 부호화된 화면의 왜곡

부호화된 화면의 왜곡은 일반적으로 양자화 계수와 참조화면과의 거리에 의해 결정된다. 일반적으로 양자화 계수가 증가하면 왜곡은 증가하고, 참조화면으로부터 거리가 증가할수록 움직임보상 이후 잔여 이미지의 분산 역시 증가한다.

움직임 보상 이후 잔여 이미지의 분산은 참조화면의 부호화 왜곡과 참조화면과의 거리와 상관이 존재하고, 다음과 같이 모델링 될 수 있다.

$$\sigma_{MC,j}^2 = D_c(q_{j_c}) + \alpha \cdot (j - j_c) \tag{7}$$

여기에서 $D_c(q_{j_c})$ 와 α는 참조화면의 부호화 왜곡과 상수이다. 첨자 j_c와 q_{j_c}는 참조화면의 화면 인덱스와 양자화 계수를 나타낸다.

현재 화면의 부호화 또는 스킵 여부를 결정하기 위해서 부호화된 화면의 왜곡 모델이 필요하다. 이전 연구에서는 참조화면의 부호화 왜곡과 참조화면과의 거리를 이용하여 현재 입력된 화면의 부호화 왜곡을 예측하였다[4].

$$D_c(q_j) = \min\{c_j \cdot q_j + d_j, \sigma_{MC,j}^2\} \tag{8}$$

여기에서 q_j, $\sigma_{MC,j}^2$, c_j, d_j는 j번째 화면의 양자화 계수, 움직임 추정후 잔여 이미지의 분산, 비트-왜곡 관련 계수들이다. 참조화면과의 거리에 의한 왜곡 변화는 잔여 이미지의 분산 값에 반영되어 있다.

주어진 양자화 계수에 대한 발생 비트량 예측을 위해 [6]에서 제안한 양자화-비트 관계를 사용한다.

$$e(q_j) = K(j)(a_j \cdot q_j^{-1} + b_j \cdot q_j^{-2}) \tag{9}$$

여기에서 $K(j)$, a_j, b_j는 비트-양자화 모델 계수이다.

3-2. 스킵된 화면의 왜곡

수신단에서는 복호시 화면 보간 등의 방법을 이용하여 스킵된 화면들의 복호 화질을 개선하는 방법들이 존재한다. 그러나, 이들은 방법에 따라 그 성능이 달라질 수 있으므로, 본 연구에서는 부호화 과정에서 스킵된 화면은 복호시 바로 직전에 복호된 화면을 보여준다고 가정한다.

선행연구에서는 실험적으로 스킵된 화면의 평균 자승오차(mean square error; MSE)는 화면간 거리에 따라 선형적으로 증가하고, 그 증가율은 인접한 참조화면들에 대해서는 거의 비슷하다는 것을 밝혔다[4]. 따라서, 본 연구에서는 스킵된 화면에 대한 왜곡은 스킵된 화면과 최근 복호된 화면 간 차이로 정의하고, 다음과 같이 [4]의 모델을 사용한다.

$$D_s(j_c,j) = D_c(q_{j_c}) + f_d(j,j_c) + \frac{(j-j_c)f_d(j_c,j_c-1)}{2}, \quad \forall i \geq j+1 \tag{10}$$

여기에서 j_c는 최근 복호된 화면의 인덱스이고, $f_d(j,i)$ 는 j번째와 i번째 화면간 차이의 MSE이다.

3-3. 채널 에러 전파 왜곡

전송 에러시에는 움직임 보상을 기반으로 하는 비디오 코덱의 특성상 화면간 에러 전파가 피할 수 없다. [7]에서는 인트라 매크로블럭 갱신률과 공간방향 루프 필터링 계수와 전송 에러간의 영향을 모델링하기 위한 이론적 체계를 확립하였다.

본 연구에서 전술한 에러 은닉 기법을 사용하였을 때 패킷 손실에 의해 화면 j에 발생하는 에러는 다음과 같다.

$$e_l(j) = \hat{f}(j_c) - \hat{f}(j) \tag{11}$$

여기에서 $\hat{f}(j)$ 는 복호기에서 j번째로 복호된 화면을 나타낸다. 에러의 분산은 다음 식과 같이 스킵된 화면의 왜곡과 동일하다.

$$\sigma_l^2(j) = D_s(j_c,j) \tag{12}$$

화면 손실에 의해 전파된 에러의 분산은 다음과 같이 계산 가능하다.

$$\sigma^2(j+k) = D_s(j_c,j)\frac{1-\beta k}{1+\gamma k}, \quad 0 \leq k < L \tag{13}$$

여기에서 $\sigma^2(j)$, γ, $\beta(=1/L)$ 는 에러의 분산, 공간방향 필터의 누출율(leakage factor), 인트라 매크로블럭 갱신률이다. 또한, L은 인트라 모드 갱신 주기이다.

그러면, 화면 $(j+k)$ 의 손실에 의해 전파되는 총 왜곡은 다음과 같이 예측 가능하다.

$$D_{ch}(j,N) = D_s(j_c,j)p_l^s(N)\sum_{k=0}^{L-1}\frac{1-\beta k}{1+\gamma k} \tag{14}$$

다중 패킷 에러는 [7]에서와 같이 여러 개의 독립된 손실들에 의한 에러를 중첩한 형태로 모델링 가능하다. 이 왜곡 모델은 패킷화 단위가 한 화면 또는 한 개 이상의 GOB를 포함하는 다른 패킷화 방식들에도 동일하게 적용 가능한데, 에러의 분산이 패킷화 단위로 계산되어야 한다는 점만 다르다.

Ⅳ. 적응형 화면 스킵 기반 소스/채널 통합 부호화

채널 적응형 비디오 부호화율 제어에 대한 기존의 연구에서는 입력 비디오 화면의 특성을 고려하면 소스 부호기의 채널 자원 감소에도 불구하고 비디오 왜곡의 증가를 완화시킬 수 있다는 결과를 도출한 바 있다[8]. 따라서, 본 절에서는 기존의 양자화 계수를 기반으로 하는 소스/채널 통합 부호화 방식에 시간방향의 부호화 계수인 화면 스킵 여부를 제어 계수로 추가한 알고리즘을 제안한다.

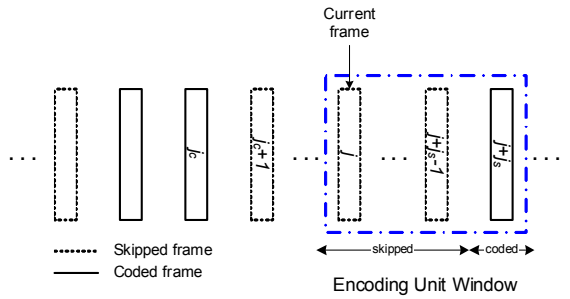

그림 2. 평균 왜곡 최소화를 위한 부호화 윈도우
Fig. 2. Encoding unit window for the optimization of average distortion.

4-1. 소스/채널 통합 부호화 윈도우

적응형 화면 스킵 기반 소스/채널 통합 부호화 방식을 위해서는 평균 왜곡을 측정하기 위한 화면 구간을 정의해야 한다. 따라서, 그림 2와 같이 현재 화면 j부터 연속 스킵될 j_s 개의 화면들과 이후 최초로 부호화되는 화면까지를 하나의 부호화 윈도우로 설정하고, 평균 화면 왜곡을 계산한다. 즉, $[j, j+j_s]$ 화면 구간을 하나의 부호화 윈도우로 설정하는데 여기에

서 $[j, j+j_s-1]$ 구간에 존재하는 j_s 개의 화면들은 스킵되고, 마지막 화면 $j+j_s$ 만 부호화된다.

실제 부호화된 화면과 스킵된 화면들을 모두 부호화된 화면으로 간주하여 부호화 왜곡을 다음과 같이 일반화 시킬 수 있다.

$$D_{enc}(j)=\begin{cases}D_c(q_j), & encoded\\ D_s(j_c,j), & skipped\end{cases} \tag{15}$$

그러면, 식 8, 10, 14를 결합하여 윈도우 내 화면들의 평균 왜곡은 다음과 같이 얻을 수 있다.

$$D_{avg}(j_s,q_{j+j_s},N)=\frac{1}{j_s+1}\sum_{i=j}^{j+j_s}D_{enc}(i)+w\cdot D_{ch}(j+j_s,N), \tag{16}$$

여기에서 w 는 평균 왜곡 결정시 소스 부호화 왜곡과 채널 왜곡의 영향을 조절하기 위한 상수로서 실험적으로 결정하였다.

4-2. 적응형 화면 스킵 기반 소스/채널 통합 부호화 문제 정의

제한된 채널 대역은 소스와 채널 부호화에 적절히 분배된다. 따라서, 채널 부호화가 없이 소스만 부호화하는 경우에 비하여 채널 부호화를 소스 부호화와 함께 수행하는 경우 소스 부호기에 할당되는 채널 자원은 감소한다. 소스/채널 통합 부호화시 소스 부호기 버퍼의 서비스율은 다음과 같이 계산된다.

$$s(j)=\frac{K}{N}R_c(j) \tag{17}$$

여기에서 $R_c(j)$ 는 j번째 화면 부호화시 가용 채널률이다.

손실 채널에서 비디오 소스/채널 통합 부호기에서 FEC 블록의 크기, 스킵되는 화면 개수, 부호화되는 화면의 양자화 계수를 조절하여 수신된 비디오의 복호시 왜곡을 최소화하는 문제는 다음과 같이 부호화 윈도우 구간별 최적화 문제로 정의될 수 있다.

$$\begin{aligned}&\min_{j_s,q_j+j_s,N} D_{avg}(j_s,q_{j+j_s},N) \\ &s.t.\quad e(j,q_j)\le\sum_{i=j}^{j+T}s(i)-B^e(j-1)\\ &\qquad p_l=p_0,\ R_c(j)=R_c\end{aligned} \tag{18}$$

여기에서 p_0 와 R_c 는 패킷 손실률과 채널률을 나타내는 상수들이나, 짧은 부호화 윈도우 구간에서만 일정한 상수로서 최적화 문제에서만 사용된다. 즉, 반복되는 최적화 문제 해법에 따라 시간에 따라 변화하는 패킷 손실률과 채널률을 고려할 수 있다.

4-3 적응형 화면 스킵 기반 소스/채널 통합 부호화를 위한 실시간 알고리즘

본 절에서는 소스/채널 통합 부호화 문제를 실시간으로 해결할 수 있는 스킵될 화면 개수, 양자화 계수, FEC 블록의 크기를 결정하기 위한 부호화율 제어 알고리즘을 제안한다.

Algorithm : Spatio-Temporal Source/Channel Rate Control

- *Step 0:* Initialization. Set $j_s^{\min}=0$, $q^{\min}=0$, $N^{\min}=\infty$, $D_{avg}^{\min}=\infty$, $B^e(0)=0$
- *Step 1:* Set the ranges of control parameters
  - Quantization parameter range: $q_{lower}=q_{prev}-\Delta q$, $q_{upper}=q_{prev}+\Delta q$
  - FEC block size range: $N_{lower}=K$, $N_{upper}=3K$
  - Number of skipped frames range: $j_s^{lower}=0$, $j_s^{upper}=J$
- *Step 2:* Estimate average distortions for possible combinations of control parameters (triple loops)
  - If $D_{avg}(j_s,q,n)<D_{avg}^{\min}$ and $e(j+j_s,q)\le\sum_{i=j}^{j+J_s+T}s(i)-B^e(j-1)$, then $D_{avg}^{\min}=D_{avg}(j_s,q,n)$, $j_s^{\min}=j_s$, $q^{\min}=q$, $N^{\min}=n$
- *Step 3:* if $j_s^{\min}>0$, skip current frame and goto *Step 5*, otherwise encode current frame with $q^{\min}$
- *Step 4:* Generate an FEC block $(N^{\min},K)$ and update $q_{prev}(=q^{\min})$
- *Step 5:* Update $B^e(j)$
- *Step 6:* Update distortion model parameters in (7), (8), (9)
- *Step 7:* Update packet loss rate(p_l) and available channel rate (R_c) based on the feedback from the receiver
- *Step 8:* Go to *Step 1*

그림 3. 적응형 화면 스킵 기반 소스/채널 통합 부호화율 제어 알고리즘

Fig. 3. Pseudo-code of the proposed joint source/channel rate control algorithm.

본 문제를 해결하기 위해 다음과 같은 알고리즘을 제안한다. 제안된 알고리즘에서는 부호화전에 입력된 화면을 분석하지 않고, 부호화한 결과를 이용하여

스킵된 화면과 부호화된 화면의 왜곡 모델과 부호화 비트 발생 모델 계수들을 갱신한다. 또한, 수신 왜곡 최소화를 위한 부호화 계수를 탐색하는 과정도 모델 기반으로 수신 왜곡을 계산하기 때문에 실시간으로 실행이 가능하다.

Ⅴ. 모의실험

본 절에서는 제안된 적응형 화면 스킵 기반 소스/채널 통합 부호화 알고리즘의 성능을 확인하기 위한 모의 실험 환경을 설명하고, 실험 결과를 제시한다.

5-1. 모의실험 환경

모의실험은 H.263+ 비디오 코덱[9, 10]을 사용하여 QCIF 비디오 시퀀스들에 대해 수행되었다. 비디오 시퀀스로는 foreman, carphone, mother&daughter, salesman을 사용하였고, 30 fps의 화면률을 갖는다. Mother&daughter는 960 frames을 나머지 시퀀스들은 300 frames을 사용하였다. 가용 채널률 R_c는 48kbps를 고정하여 사용하였다. 다양한 비디오 서비스 유형과 손실 채널 환경에 대해 제안된 알고리즘의 성능을 파악하기 위해서 다음 표와 같이 종단간 지연, 인트라 매크로블럭 갱신률, 패킷 손실률의 제어 변수 범위를 설정하고 각각의 조합에 대해 모의실험을 수행하였다.

표 1. 모의실험 환경 변수
Table 1. Simulator Parameters

	Value	Units
End-to-end delay(T)	5, 10, 20	Frame
Intra MB update rate(L)	5, 10, 15, 20, 25, 30, 60	Frame
Packet loss rate(p_l)	0~60 with step 4	Percentage(%)

제안방식의 성능은 TMN8[10] 부호화율 제어 방식과 비교되었다. 성능 지표로서 비디오 시퀀스 전체에 대한 평균 PSNR, 주관적 화질이 사용되었다. 스킵된 화면들에 대해서 직전에 복호된 화면을 복호된 화면으로 간주하여 PSNR을 계산하였다. 기언급된 바와 같이 스킵된 화면의 PSNR은 화면 보간이나 에러 은닉 기법등을 사용하면 향상될 수 있으나, 본 연구에서는 고려하지 않는다. 양자화 계수의 최대 변이 Δq는 3으로 설정되었는데, 이는 이전 화면과 현재 부호화되는 화면의 양자화 계수의 큰 차이에 의한 급격한 화질 변이를 방지하기 위함이다.

5-2. 비디오 시퀀스별 PSNR 성능

적응형 화면 스킵 기반 소스/채널 통합 부호화 알고리즘의 성능을 확인하기 위해 패킷 손실이 존재하는 상황에서 TMN8 부호화율 제어 방식과 화면별 PSNR을 비교하였다. 그림 4는 20%의 패킷 손실이 존재하는 채널에서 수신단에서의 복호 화질을 나타낸다. 제안방식에서 스킵된 화면을 빼고 실제 부호화된 화면의 화질만 표시한 경우(Coded)와 스킵된 화면까지 포함하여 화질을 표시한 경우(Coded+Skip) 두 가지를 표시하였다. 모의실험 결과 각 표현 방식 모두 TMN8 방식에 비해 높은 PSNR 성능을 얻을 수 있음을 확인할 수 있다. 제안방식은 채널 상황 및 입력 비디오의 왜곡 특성을 고려하여 적절히 화면을 스킵하고, FEC 블록 크기를 결정하므로 부호화에 사용될 비트 자원을 절약할 수 있어서 다수의 패킷 손실이 발생하는 구간에도 매우 우수한 PSNR 성능을 보이고 있다.

5-3. 패킷 손실률에 따른 평균 PSNR 성능

제안방식의 패킷 손실률에 따른 성능을 비교하기 위해 패킷 손실률을 0%에서 60%까지 4%씩 증가시키면서 평균 PSNR 성능을 TMN8과 비교하였다. 그림 5에서와 같이 모든 비디오 시퀀스들에 대해 제안방식과 TMN8 모두 패킷 손실률이 증가함에 따라 평균 PSNR은 감소하는 경향을 보였다. 그러나, TMN8 방식은 패킷 손실률의 증가에 따라 전반적으로 동일한 기울기로 PSNR이 감소하는 결과를 확인하였으나, 제안방식의 경우 약 40~50%의 패킷 손실률까지의 PSNR 저하는 그리 크지 않고, 그 이상의 패킷 손실률 구간에는 급격한 PSNR 저하를 확인할 수 있다. 이 현상은 특정 패킷 손실률 범위 안에서 제안방식이 수신 왜곡을 최소화하는 방향으로 화면 스킵과 FEC 블록 크기를 제어하고, 화면 스킵에 의해 절약된 비

트량을 FEC 블록 크기 증가에 효과적으로 사용할 수 있었기 때문으로 설명할 수 있다. 그러나, 특정 패킷 손실률을 넘어가는 경우에는 패킷 손실을 막기 위해서는 FEC 블록의 크기가 상당히 커져야 하는데, 이에 소모되는 비트량을 과도한 화면 스킵으로 해결하거나, 또는 FEC 블록의 크기를 요구되는 만큼 키울 수 없기 때문에 급격한 화질 저하가 발생된다.

Mother&daughter와 salesman 시퀀스에 대해 제안방식은 패킷 손실률이 40% 이하일 때 TMN8에 비하여 약 0.8~2.0 dB의 PSNR 개선을 확인할 수 있었다.

인트라 매크로블럭 갱신주기를 15, 30, 60으로 변화시켰을 경우에도 mother&daughter 시퀀스에 대해 제안방식이 TMN8과 비교하였을 때 평균 1 dB의 평균 PSNR 이득이 있음을 확인할 수 있었다.

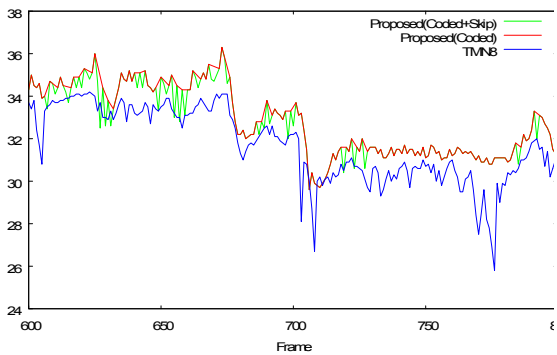

그림 4. 제안방식의 PSNR 성능, T=5, L=30, PLR=0.2, mother&daughter

Fig. 4. PSNR comparison: proposed vs. TMN8.

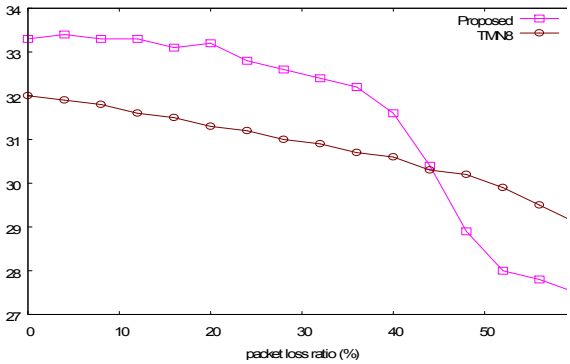

그림 5. 패킷 손실률에 따른 평균 PSNR 비교, T=5, L=30, salesman

Fig. 5. Average PSNR over packet loss rate.

5-4. 주관적 화질 비교

실제 비디오 사용자들이 느끼는 화질을 비교하기 위해 carphone과 mother&daughter 시퀀스들에 대해 제안방식과 TMN8 방식의 주관적 화질 비교를 수행하였다. 패킷 손실률이 20%인 채널 상황에서 인트라 매크로블럭 갱신주기 30을 사용한 경우에 대한 수신단에서 복호된 화면을 그림 6에 나타내었다. 테스트한 비디오 시퀀스들에 대해서 모두 제안방식이 TMN8 방식에 비해 비디오 시퀀스내 인물의 얼굴 윤곽이 뚜렷하며 패킷 손실에 따른 화질 열화가 덜 함을 확인할 수 있다.

VI. 결 론

본 연구에서는 손실 채널에서 비디오의 수신 왜곡을 최소화하기 위해 소스/채널 통합 부호화율 제어 알고리즘을 제안하였다.

(a) 57번째 화면

(b) 203번째 화면

그림 6. 주관적 화질 비교, carphone

Fig. 6. Subjective quality comparison, carphone

제안된 알고리즘은 비디오 전송시스템의 버퍼 제약을 기반으로 화면 스킵, 양자화 계수, FEC 블록 크기를 제어계수로 하는 소스/채널 통합 부호화율 제어 구조이다. 제한된 채널률과 패킷 손실률에 대한 비디오 수신단의 평균 왜곡 최소화 문제를 정의하고, 실시간 해결을 위한 부호화 윈도우 기반 알고리즘을 제

안하였다. 모의실험을 통해 제안방식은 40% 이하의 패킷 손실률과 30 이상의 인트라 매크로블럭 갱신률에 대하여 TMN8보다 우수한 PSNR과 주관적 화질 성능을 얻을 수 있었다.

제안된 알고리즘은 채널 손실이 존재하는 실시간 비디오 통신 환경에서 채널 오류에 강인하면서도 우수한 수신 화질이 요구되는 고화질 영상통신 및 방송 시스템 등에 적용될 수 있다.

감사의 글

본 논문은 2006년 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원(KRF-2006-311-D00699)과 2007년도 한국항공대학교 교비지원 연구비에 의하여 지원된 연구의 결과임

참 고 문 헌

[1] A. Vetro, Y. Wang, and H. Sun, “Estimating distortion of coded and non-coded frames for frameskip-optimized video coding,” *IEEE ICME,* pp. 541-544, 2001.

[2] S. Liu and C.-C. J. Kuo, “Joint temporal-spatial bit allocation for video coding with dependency,” *IEEE Trans. Circuit Syst. Video Technol.,* Vol. 15, No. 1, pp. 15-26, 2005.

[3] R. C. Reed and J. S. Lim, “Optimal multidimensional bit-rate control for video communication,” *IEEE Trans. Image Processing,* Vol. 11, No. 8, pp. 873-885, 2002.

[4] M. Lee and D. Lee, “Delay constrained spatio-temporal video rate control for time-varying rate channels," *PCM 2006, LNCS 4261*, pp. 537-546, Nov. 2006.

[5] S.-H. Gary Chan, et. al., ”Video loss recovery with FEC and stream replication,” *IEEE Trans. Multimedia,* Vol. 8, No. 2, pp.370-381, Apr. 2006

[6] T. Chiang and Y.-Q. Zhang, “A new rate control scheme using quadratic rate distortion model,” *IEEE Trans.* Circuit Syst. Video Technol., Vol. 7, No. 1, pp. 246-250, 1997.

[7] K. Stuhlmuller, N. Faber, M. Link, and B. Girod, “Analysis of video transmission over lossy channels,” *IEEE J. Selected Areas in Commun.,* Vol. 18, No. 6, pp. 1012-1032, 2000.

[8] M. Lee and J. Kim, “Adaptive video frame rate control over explicit rate networks based on channel rate prediction,” *IEICE Trans. Communi.*, Vol. E84-B, No.7, pp.1915~1925, July 2001.

[9] ITU-T Recommendation H.263, Version 2, 1998.

[10] ITU Study Group 16, Video Coding Experts Group, Video Codec Test Model, Near-Term, Version 10 Draft 1, 1998.

이 명 진 (李明辰)

1994년 2월 : KAIST 전기및전자공학과(공학사)
1996년 2월 : KAIST 전기및전자공학과(공학석사)
2001년 8월 : KAIST 전자전산학과(공학박사)
2001년 3월 ~ 2004년 2월: 삼성전자 SYSTEM LSI 책임
2004년 3월 ~ 2007년 2월: 경성대학교 전기전자공학과 조교수
2007년 3월~ 현재: 한국항공대학교 조교수
관심분야 : 멀티미디어 통신, 비디오 코딩, 컴퓨터 비전