

# 인터넷 환경에서 FEC 기능이 추가된 AMR 음성 부호화기를 이용한 오디오 스트림 전송

학생회원 김은중\*, 정회원 이인성\*\*

## Audio Stream Delivery Using AMR(Adaptive Multi-Rate) Coder with Forward Error Correction in the Internet

Eunjoong Kim\* *Student Member*, Insung Lee\*\* *Regular Member*

### 요 약

본 논문에서는 ETSI와 3GPP에서 차세대 이동통신 IMT-2000 서비스의 음성부호화기의 표준으로 채택한 AMR을 인터넷을 통한 멀티미디어 서비스에서 사용하기 위해 부가 정보를 이용한 손실 패킷 복구 방법이 첨가된 전송 방법을 제시한다. 인터넷과 같은 패킷 교환 망에서의 음성 통신에서 과도한 패킷 손실은 급격한 음질 저하를 유발한다. 본 논문에서는 음성 패킷 데이터를 순방향 오류정정(FEC)의 부가 정보로 사용하고 연속 패킷 손실이 발생하였을 경우 오류 은닉방법을 사용하여 패킷 손실에 의한 음질 저하를 개선하는 방법을 제안한다. 순방향 오류정정 방법 중 부가 음성 정보를 원래의 음성정보와 함께 보냄으로써 손실된 음성은 부가 음성 정보를 이용해 복구할 수 있다. 본 연구에서 사용한 AMR 음성 부호화기는 CELP기반의 음성 부호화기 이므로 음성 부호화기의 특징을 이용해 2개 이상의 군집오류가 발생했을 경우 패킷 손실이 일어나기 전후의 데이터를 이용해서 손실된 패킷으로 인한 영향을 최소로 하는 오류은닉 방법을 사용하였다. 제안된 방법의 성능을 평가하기 위해 AMR 부호화기의 고음질 압축 방법인 12.2 kbit/s 모드로 전송하는 방법과 ITU-T 표준안인 CS-ACELP로 전송하는 방법을 SNR과 MOS 측정을 통해 비교하였다. 제안된 방법이 10%의 평균 패킷 손실률에서 부호화기 자체의 오류은닉 기술을 적용한 AMR - 12.2 kbit/s 모드보다 MOS값에서는 1.1, SNR값은 5.61 dB 높았으며, 제안된 방법은 20%의 손실률에서도 통신 가능한 음질을 유지하였다.

### ABSTRACT

In this paper, we present an audio stream delivery using the AMR (Adaptive Multi-Rate) coder that was adopted by ETSI and 3GPP as a standard vocoder for next generation IMT-2000 service in which includes combined sender (FEC) and receiver reconstruction technique in the Internet. By use of the media-specific FEC scheme, the possibility to recover lost packets can be much increased due to the addition of repair data to a main data stream, by which the contents of lost packets can be recovered. The AMR codec is based on the code-excited linear predictive (CELP) coding model. So we use a frame erasure concealment for CELP-based coders. The proposed scheme is evaluated with ITU-T G.729 (CS-ACELP) coder and AMR - 12.2 kbit/s through the SNR (Signal to Noise Ratio) and the MOS (Mean Opinion Score) test. The proposed scheme provides 1.1 higher in Mean Opinion Score value and 5.61 dB higher than AMR - 12.2 kbit/s in terms of SNR in 10% packet loss, and maintains the communicable quality speech at frame erasure rates up to 20%.

\* 충북대학교 전기공학과 통신신호처리연구실(nirvana@tsp.chungbuk.ac.kr),

\*\* 충북대학교 전기공학과 통신신호처리연구실(inslee@cbucc.chungbuk.ac.kr)

논문번호 : K01179-0810, 접수일자 : 2000년 8월 10일

※ 본 연구는 한국과학재단 특정기초 연구지원에 의해 연구되어짐(과제번호 2001-1-30200-003-3)

## I. 서론

현재 통신 서비스의 발전 및 연구 방향은 광대역 멀티미디어 서비스의 통합에 집중되어있다. 인터넷의 상용화와 Web의 등장은 인터넷 사용자의 급속한 증가와 더불어 이용분야의 확대를 초래했다. 그 예가 바로 무선 네트워크 환경에서의 멀티미디어 응용인데 이것은 인터넷의 성공에 의해 대중화된 것이다. 이동 단말기를 통한 인터넷 서비스는 인터넷을 통한 통신서비스가 혁명적으로 변화할 수 있는 가능성을 보여주는 예인데, 이런 무선 환경에서의 서비스를 지원하고 확대시키기 위해 IMT-2000 시스템이 제안되었고 개발되었다.

인터넷의 팽목할 만한 성장과 전 세계적인 데이터 전송능력의 증가는 인터넷을 통한 실시간 음성 통신이 기존의 통신 매체를 대체할 대안 통신 매체로서의 가능성을 열어 놓은 계기가 되었다. 그러나 인터넷은 패킷 교환방식 네트워크의 특징인 높은 전달 지연 및 패킷 손실 등으로 인하여 실시간 음성 통신에는 적합하지 않다. 이런 환경의 인터넷에서 실시간 멀티미디어 통신 서비스를 가능케 하는 핵심 기술에는 음성 코딩 및 압축기술이 있는데, 이 기술에 따라 데이터의 양과 음질에 커다란 차이가 있기 때문에 매우 중요한 기술이다. 그리고 현재 RTP(Real-time Transport Protocol)가 표준으로 채택된 인터넷 상의 실시간 데이터 전송 프로토콜의 구현 및 데이터 전송에서 나타나는 패킷 손실에 대한 실시간 복구 기술이 있다.<sup>[1]</sup> 이 기술은 트래픽이 심할 경우 40% 까지 나타나는 패킷손실, 특히 수개 이상 연속적으로 발생하는 패킷 손실을 고려할 경우 인터넷을 통한 멀티미디어 서비스의 품질을 좌우하는 중요한 기술이다.

인터넷 상에서 오디오 정보를 전송할 때 나타나는 패킷 손실은 매우 다양하여 전송패킷의 종단간 지연(End-to-End Delay), 지터(Jitter), 손실(Packet Loss), 비순차 패킷(Out-of-Order Packet)등이 발생하게 되며, 이러한 패킷손실의 처리 알고리즘에 따라 통신 품질에 큰 영향을 미치는 것으로 알려졌다.<sup>[2][3][4]</sup> 그러므로 인터넷 상에서 실시간 데이터 서비스를 제공해야 하는 응용의 경우 인터넷이라는 공유망을 사용함으로써 불가피하게 발생하는 손실 패킷을 제어 혹은 복구하는 기술이 매우 중요하다.

본 논문에서 제시한 방법은 AMR 음성 부호화기의 특징을 이용한 부가 정보를 이용한 손실 패킷

복구 방법 및 오류 은닉 방법을 제안하였다. 여러 모드로 동작하는 부호화기의 특징을 이용하여 7.40 kbit/s 모드를 기본으로 하고 4.75 kbit/s 모드를 부가 정보로 하여 전송해서 패킷 손실의 대부분을 차지하는 단일 패킷 손실 복구에 사용했다. 부가 오디오 정보로 복원이 불가능한 손실 패킷에 대해서는 CELP 기반 부호화기인 AMR 음성 부호화기의 특징을 이용하여 대체 패킷을 생산해내는 방법과 예측 부호화기의 특징을 이용해 상태 오류(State Error)를 줄이는 오류 은닉방법을 사용했다.

본 논문 II장은 AMR 음성 부호화기 구조에 대해서 설명하고 III 장은 음성 복원 및 은닉 기술(Voice Reconstruction and Error Concealment Techniques)에 대하여 설명하고, IV장에서는 실험결과 및 객관적 음질 평가 결과인 SNR값 및 주관적 음질 평가 결과를 다룬다.

## II. AMR 음성부호화기

AMR음성부호화기는 ACELP(Algebraic Codebook-Excited Linear Prediction)에 기반을 두고 있다. 또한 전송률을 여러 개 가짐으로써 MR-ACELP(Multi-Rate ACELP)라고 하기도 한다. AMR 음성 부호화기는 무선 채널의 환경변화에 적응적으로 대처하며, 명료한 음질(toll quality)을 가진다는 장점이 있다. 이에 ETSI에서는 전송률을 4.75 kbit/s에서 12.2 kbit/s까지 8개의 모드를 가지는 하나의 통합된 음성부호화기를 구성했다. 입력되는 음성 신호는 8 kHz로 표본화되며 한 개의 프레임 길이가 5 ms인 4개의 부프레임을 갖는 20 ms의 프레임을 갖는다. 모드에 따른 전송률과 각 모드의 파라미터 개수는 <표 1>과 같다.

표 1. AMR 부호화기의 모드별 비트율/파라미터 개수

Codec mode	Source codec bit rate	par. no
AMR_12.20	12.20 kbit/s(GSM-EFR)	244
AMR_10.20	10.20 kbit/s	204
AMR_7.95	7.95 kbit/s	159
AMR_7.40	7.40 kbit/s(IS-641)	148
AMR_6.70	6.70 kbit/s(PDC-EFR)	134
AMR_5.90	5.90 kbit/s	118
AMR_5.15	5.15 kbit/s	103
AMR_4.75	4.75 kbit/s	95
AMR_SID	1.80 kbit/s	

GSM-EFR : ETSI GSM06.90 Enhanced Full Rate  
 IS-641 : TTA/EIA IS-641 TDMA Enhanced Full Rate  
 PDC-EFR : 6 ARIB 6.7 kbit/s Enhanced Full Rate

AMR 음성부호화기는 비 음성 구간에서 전송률을 조정하여 효과적인 전송과 전송용량을 증가시키기 위해서 SCR(Source Controlled Rate)부분과, 음성과 비음성에 관한 정보를 나타내는 VAD(Voice Activity Detector)부분, 비 음성 구간에 삽입하는 CNI(Comfort Noise Insertion) 부분, 무선 환경에서 발생하는 전송 오류나 손실프레임의 영향을 줄이기 위한 오류 은닉(Error Concealment) 부분을 가지고 있다.

일반적으로 CELP(Code-Excited Linear Prediction)부호화 방식은 크게 LP(Linear Prediction)분석과 적응 코드북 탐색(adaptive codebook)과 고정 코드북 탐색(fixed codebook search)으로 구성되어 있다. AMR 인코더의 구조는 <그림 1>과 같다.

입력 음성은 직류성분 제거 등을 위한 고역통과 필터 등의 전처리 과정을 거친 후 12.2 kbit/s에서는 30ms의 길이를 가지는 서로 다른 두 개의 비대칭 창함수를 사용하고, 그 이외의 모드에서는 한개의 비대칭 창함수를 사용하여 10차의 LP분석을 수행하게 된다. 이때 12.2 kbit/s에서는 미래의 샘플값을 참조하지 않고 단순히 현재의 프레임 값을 사용하고, 나머지 모드에서는 5ms의 미래 샘플값을 참조한다. 계산된 LP filter 계수는 양자화와 보간을 위하여 LSP(Line Spectrum Pair) 파라미터로 변환되는데 12.2kbit/s의 경우 두 set의 LP 파라미터들은 두 개의 1 차 MA(Moving Average) 예측을 사용하여 두 개의 LSF residual 벡터를 구하게 되고 이는  $2 \times 2$  영역으로 구성된 5개의 submatrices를 가지는 SMQ(Split Matrix Quantization)을 사용해서 양자화 하게 된다. 나머지 모드에서 LSF residual 벡터는 3개의 subvector로 나누고 각각 7 ~ 9 bit를 사용해서 양자화 한다.

적응 코드북 검색(Adaptive codebook search)을 위한 피치 예측(pitch prediction) 과정은 고정된 lag 영역에서 개구간 검색(open-loop search)과 폐구간 검색(closed-loop search)을 수행함으로써 이루어진다. weighting filter를 거친 원음성 신호로부터 최적  $T_0$ 의 정수 lag,  $T_0$  값을 구하기 위해 개구간 검색을 하게 된다. 이렇게 구해진  $T_0$  값을 이용해서 폐구간 검색과정을 수행하게 된다. 4.75 kbit/s, 5.15 kbit/s에서는 매 프레임당 한번 개구간 검색을 시행하고 나머지 모드에서는 두 번 시행한다. 폐 구간검색은 일반적으로 계산범위와 분해능에 차이를 보이는데, 12.2 kbit/s에서는 1/6, 나머지 모드에서는 1/3의

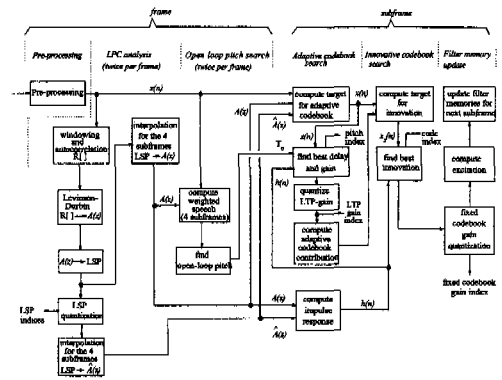


그림 1. AMR 인코더의 구조

분해능을 가지고 구한다.

최적 피치 지연 값을 찾아내면 기중 필터의 입력 음성신호와 가중 합성 음성신호 사이의 MSE(Mean Square Error)를 최소화 시키는 값을 찾는 고정코드북 검색(Algebraic codebook search)이 이루어진다. 고정 코드북 검색을 위한 목적(target) 신호는 선형 예측 잔여 신호에서 적응 코드북의 기여분을 제거함으로써 얻어진다. 고정 코드북은 모드별 비트 전송률의 주된 변화를 보이는 부분인데, 변화하는 비트 전송률의 범위는 7.0 kbit/s에서 1.8 kbit/s 이다. 모드에 따라 코드북의 차이를 보인다. 하나의 서브프레임을 구성하는 40개의 펄스의 위치는 2~5개의 트랙으로 구분되어진다. 각 부 프레임에서 모드별 2개에서 10개까지의 펄스가 선택되고 그 펄스들의 위치와 크기를 부호화하는데, 이 펄스 수의 변화로 비트 전송률이 변하게 된다. 펄스의 크기는 +1~-1의 범위를 가지며 부호 비트와 위치를 나타내는 비트를 전송하게 된다.

오류 은닉(Error Concealment)의 목적은 프레임 손실에 의한 음질 저하를 줄이는 것이다. 상태기구(state-machine)를 이용하는 오류 은닉 알고리즘은 복호화기에서 사용된다. AMR의 오류 은닉은 크게 두 가지 경우를 고려한다. 현재 프레임이 손상되었을 경우와 현재 프레임은 올바르게 수신되었지만 이전 프레임이 손상되었을 경우 다르게 오류 은닉을 적용한다. 현재 프레임은 손상되지 않았지만, 이전 프레임이 손상되었을 경우는 LTP 이득과 고정 코드북의 이득을 이전에 마지막으로 올바르게 수신된 프레임의 파라미터 값 이하로 제한하고, 현재 프레임이 손실된 상황에서는 LTP 이득과 고정 코드북 이득을 이전 프레임의 값을 감쇄시켜 사용하는 오류 은닉방법을 사용한다.

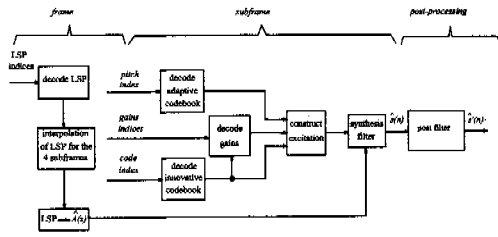


그림 2. AMR 디코더 구조

복호화 과정에서는 전송된 파라미터 값들로부터 LP 파라미터, 피치 코드북 인덱스 그리고 이득 값들을 추출하며 이를 이용하여 합성음을 복원하게 된다. 또한, 복원된 합성음의 음질을 향상시키기 위해 post-filtering을 통과시킨다. AMR의 decoder의 구조는 <그림 2>와 같다.

### III. 음성복원 및 오류 은닉 방법

#### 3.1 손실패킷 복원 및 오류은닉 방법

효율적인 손실패킷 복구를 위해서는 패킷의 손실 특성에 맞게 복구를 하는 것이 중요하다. Mbone (Multicast Backbone)에서의 패킷손실 특성에 관한 연구가 많이 진행되어 왔는데<sup>[4][6][7]</sup>, 많은 수의 node에 패킷을 전송할 때 수신 단에서의 패킷손실은 피할 수 없는 사실임이 증명되었다. 그러므로 인터넷을 통한 음성 전송에서 손실 패킷에 대한 복원 및 오류 은닉은 중요한 문제이다.

손실 패킷 처리 방법에는 여러 가지가 제시되었는데, 송신자 기반 손실 패킷 복구방법과 수신단에서 독립적으로 이루어지는 손실 은닉 기법 등이 있다. 송신자 기반 복구 방법을 크게 두 가지로 나누면 재전송을 통해 복구하는 능동적인 방법과 채널 코딩을 이용하는 수동적인 방법이 있다. 수동적인 채널 코딩은 매체 독립적인 FEC와 매체 종속적인 FEC를 포함하는 FEC를 이용하는 방법과 인터리빙 기반 기법들로 나누어진다.

오류 은닉 방법은 오디오 스트림의 수신자에 의해 시작되고, 송신자로부터의 어떤 도움도 필요로 하지 않는 방법으로 손실 패킷을 대체할 원본과 비슷한 패킷을 만들어 내는 방법이다. 이 방법이 가능한 것은 오디오 신호가 단구간 자기 유사(short-term self similarity) 특징이 많은 부분에서 나타나기 때문이다. 그래서 이 방법은 비교적 적은 손실률(15% 이하)과 작은 패킷 (4-40ms)크기 상에서 동작한다.

이 방법에는 손실 부분을 묵음이나 잡음 또는 이전 패킷의 값으로 채워 넣는 삽입 기반(Inserting Based)의 오류 은닉 방법과 손실 패킷과 비슷한 대체 패킷을 패턴 매칭이나 보간법을 이용해 생성해 내는 방법인 보간법 기반(Interpolation Based) 오류 은닉 방법과 손실된 패킷 주위에서 디코더의 상태를 알아내고 그것을 통해 손실된 패킷을 대체할 패킷을 다시 만드는 방법인 재생산 기반(Regeneration Based) 오류 은닉 기법이 있다.<sup>[5]</sup>

#### 3.2 제안된 시스템

본 논문에서 제안한 손실패킷 복구 및 오류 은닉 방법이 첨가된 시스템의 구조는 <그림 3>과 같다. 송신 단에서는 부가 정보를 이용한 손실 패킷 복원 방법 적용을 위해 입력된 음성을 고 비트율과 저 비트율로 두 번 부호화하고 저 비트율로 부호화한 정보를 다음 패킷의 부가 정보로 사용하기 위해 한 프레임의 지연 시킨다. 현재 프레임에서 고 비트율로 부호화해 생성된 정보와 이전 프레임에서 생성되어 지연된 정보를 하나로 합쳐 한 개의 패킷으로 구성을 한다. 이렇게 생성된 패킷은 수신측에서 손실된 패킷의 복구 혹은 오류 은닉방법에 사용된다.

수신 측에서는 수신된 패킷의 손실 여부를 검사하여 단일 패킷(n+1번째 패킷)이 손실되었을 경우에는 다음 패킷(n+2번째 패킷)에 첨가되어 있는 부가 정보로 손실된 패킷을 복원하고 2개의 연속된 패킷이 손실되었을 경우에는 우선 다음 패킷(n+2번째 패킷)의 부가 정보로 한프레임(n+1번째 패킷)을 복구하고, 이전에 올바르게 수신된 패킷(n-1번째 패킷)과 부가 정보로 복원된 패킷(n+1번째 패킷)의 보간을 통하여 부가 정보로는 복구가 불가능한 패킷(n번째 패킷)을 재생산한다. 만약 2개 이상의 패킷이 손실되게 되면(그림 3에서 n-1번째 패킷도 손실된 경

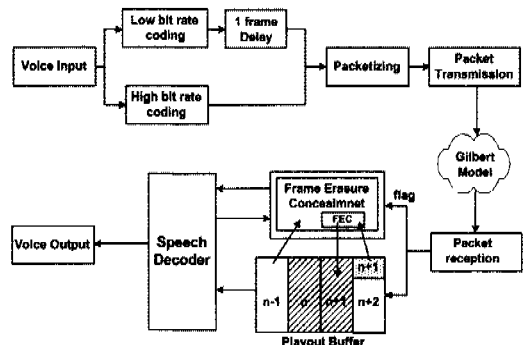


그림 3. 제안된 시스템 구조

우) 이전에 올바르게 수신된 패킷(n-2번째 패킷)을 복사하는 방법을 사용하였는데, 이때 CELP기반의 부호화기 특징인 상태의 비동기로 인한 오류의 전파를 줄이기 위하여 과거 여기 신호를 초기화하는 방법을 적용하여 음질 저하를 줄였다.

### 3.3 부가 정보를 이용한 손실 패킷 복원

부가 음성 정보를 이용하는 매체 종속 FEC방법은 <그림 4>와 같이 음성 정보를 두 가지 이상 압축방법으로 중복 부호화하여 서로 다른 시간대에 전송하는 방법을 말한다. 일반적으로 고 음질, 고 비트율을 갖는 음성코딩(ADPCM, GSM)과 저 음질, 저 비트율의 음성 코딩(LPC)을 동시에 행하면서, 서로 다른 시간에 코딩된 오디오 정보를 하나로 패킷화하여 전송하고 수신단에서는 저 비트율 코딩 정보를 이용하여 손실 패킷을 복원하는 것이다. 이 방법은 Mbone상에서 사용하기 위하여 Hardman등<sup>[3]</sup>과 Bolot 등<sup>[4]</sup>에 의해 제안되었으며, Podolsky등<sup>[2]</sup>에 의해 실험되었다.

이와 같은 과정에 의하여 생성된 패킷을 수신측에서 수신하였을 때 만일 패킷이 <그림 5>와 같이 손실되었다면 손실된 패킷 후에 수신될 패킷에 포함되어 있는 저비트율의 부가 오디오 정보를 이용하여 <그림 5>에서와 같이 복구하게 된다.

본 논문에서는 AMR 부호화기의 특징인 다(多) 전송률을 이용했다. 즉 두개의 다른 음성 부호화기를 사용한 것이 아니라, 여러 모드가 존재하는 AMR 음성 부호화기만을 사용해 고 비트율 정보와 저 비트율 정보를 만들어 냈다. 송신 단에서는 한 프레임에 대해 7.40 kbit/s 모드의 고 비트율 압축 방법과 4.75 kbit/s 모드의 저 비트율 압축 방법을 통해 원 음성 정보와 부가 음성 정보를 생성해서, 이전에 만들어진 부가 음성 정보와 원 음성 정보를

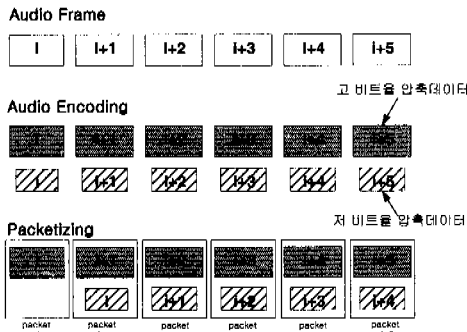


그림 4. 음성 패킷의 부가 정보 추가

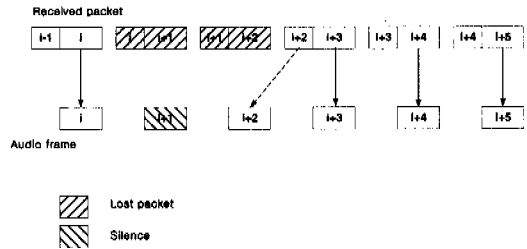


그림 5. 부가 오디오 정보를 이용한 손실 패킷 복원

함께 전송하는 FEC 방법을 사용했고, <그림 5>에 나타난 것과 같이 수신 단에서는 정상적으로 패킷을 수신하면 원래의 고 비트율 정보를 이용해 복호화하고 손실을 발견하면 다음 패킷에 첨가되어 전송되어진 손실 패킷에 대한 부가 음성 정보로 손실된 패킷을 복원하였다. 부가 오디오 정보를 이용하는 매체 종속 FEC방법은 손실 패킷 복구에 가장 널리 사용되는 방법으로 약간의 대역폭 증가로 많은 이득을 얻을 수 있다<sup>[5]</sup>. 그러나 두개 이상의 패킷 손실시 복구가 불가능한 단점이 존재한다. 현재 인터넷 환경에서 패킷 손실이 간헐적으로 단일 패킷만 발생하는 것이 아니며 연속된 패킷 손실도 존재하므로 연속된 패킷 손실에 대한 대책이 없으면 서비스 품질에 큰 저하가 발생할 수 있다.

### 3.4 군집오류 발생시 오류 은닉 방법

부가 오디오 정보를 이용한 손실 패킷 복원방법을 사용하였을 경우 부가 오디오 정보로도 복원할 수 없는 연속 패킷 손실이 <그림 5>에서와 같이 발생하게 되었을 경우 기존의 오류 은닉 방법으로는 손실된 프레임에 묵음이나 백색잡음으로 대체하거나, 이전에 수신된 패킷을 복사하는 방법이 사용되었으나, 본 논문에서는 음성 부호화기의 특징을 이용해 복원 불가능한 패킷을 재생산하는 방법으로 오류를 은닉하였다.

음성 부호화기 중 현재 최고의 손실 프레임 복원 방법은 이전에 올바르게 수신한 프레임의 파라미터를 감쇄시켜서 손실된 프레임에서 사용하는 방법이다. 이 방법은 부가적인 지연을 필요로 하지 않는 장점이 있기 때문에 최근의 음성 부호화기 표준에 사용되었다. 그러나 일률적으로 파라미터 값을 감쇄시키기 때문에 연속적인 프레임 손실 발생시 음질 왜곡이 심각하게 발생한다.

본 논문에서 사용한 오류 은닉 방법은 부가 정보를 이용한 손실 패킷 복원 방법으로 복구 불가능한 연속 패킷 손실 발생시 적용 하였다. 즉 2개의 연속

패킷이 손실되었을 경우 한 개의 패킷은 부가 오디 오 정보로 복원이 가능하지만 한개는 복원이 불가능하다. 이 경우 복원이 불가능한 패킷을 기준으로 부가 정보로 복원된 패킷은 미래 패킷이 되고, 이전의 올바르게 수신된 패킷은 과거 패킷이 된다. 복원 불가능한 패킷을 미래 패킷과 과거 패킷의 파라미터들 간의 보간법을 적용하여 대체 패킷을 생산하는 방법을 사용하였다. 여기서 적응 코드북 이득은 선형 보간법을 적용하였고 적응 코드북 인덱스는 선형 보간법 대신에 미디어언 평활(median smoothing) 방법을 사용하였다. 그러나 LSF와 고정 코드북 이득은 보간법을 사용하지 않고 미래 패킷의 파라미터를 사용하였다. 3개 이상의 연속 패킷 손실이 발생하면 이전 마지막으로 올바르게 수신한 패킷을 단순히 복사하여 사용하는 오류 은닉 방법을 사용하였다.

3.5 상태 오류(State Error) 축소

CELP 기반 부호화기는 음성의 현재 상태(state)를 기반으로 한 모델 파라미터를 생성한다. 이 파라미터들을 전송하고 복호화기에서는 이 파라미터로 자체 상태를 생성해서 음성을 재생한다. 일반적인 작동에서는 부호화기와 복호화기 양쪽 상태가 같이 갱신되므로 복호화기에서는 전송된 파라미터로부터 원음을 재생할 수 있다. 그러나 패킷 손실이 발생하면 부호화기와 복호화기의 상태 동기가 더 이상 맞지 않으므로 복호화기에서는 원활히 상태를 갱신할 수 없게 된다. 그런 까닭에 더 이상 패킷 손실이 발생하지 않는다 해도 복호화기에서는 정확히 갱신되지 않은 상태를 이용해 올바르게 수신된 패킷에 대해서도 복호화기 때문에 음질저하가 발생한다.

<그림 6>은 AMR - 7.4 kbit/s 모드에서 2개의 연속 패킷 손실(1761 - 2080 samples)이 발생했을 경우와 패킷 손실이 발생하지 않았을 경우와의 SE(Square Error)를 나타낸 것이다. 연속 패킷 손실로 패킷 손실이 발생한 구간과(1761-2080 samples) 패킷 손실 이후의 구간(2081-3040 samples)에서 오류가 발생한 것을 알 수 있다. 패킷 손실 이후의 오류를 상태오류라고 하는데<sup>[13]</sup>, 상태오류의 크기는 패킷 손실 때와 같고 오래 지속된다는 것을 알 수 있다. CELP 기반 부호화기는 예측 부호화기 이므로 각 프레임은 현재 프레임에 의해 생성된 정보 비트 뿐만 아니라 이전 프레임으로부터 영향을 받는 예측 상태를 기반으로 복호한다. 그러므로 부정확한 예측 상태는 손실된 프레임 이후의 신호를 복호하

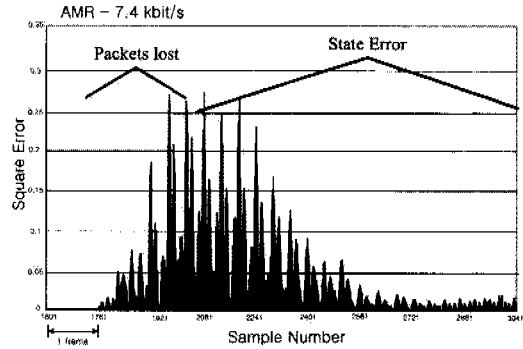


그림 6. 군집 오류시 SE(Square Error)

는데 음질 저하를 유발시키므로 예측 상태를 정확히 갱신하지 않고 음성신호를 복원해내서는 안 된다. 본 논문에서 사용하는 오류 은닉방법은 복호화기에서 복원하지 못하는 프레임에 대해 과거 여기 신호를 '0'으로 초기화 하는 방법을 사용하여 상태 오류를 줄였다.

IV. 시뮬레이션 및 성능평가

4.1 실험환경

실제 인터넷 환경에서 대다수의 수신 단들은 2 ~ 5% 가량의 패킷손실을 겪고 있는데 패킷의 손실은 대부분 하나씩 일어나지만 순수한 균일 분포보다는 연이어서 일어나는 손실이 많다. 그러므로 오류를 복구하는데 어떤 오류 제어 기술을 사용하는가를 결정하는 과정에서 오디오 패킷의 손실 특성은 중요한 역할을 한다. 오디오 패킷 손실과정의 상관구조를 Gilbert 모델과 같은 저 차의 마르코프 체인(Markov chain)으로 나타낼 수 있다는 연구결과가 나왔다. [4][8][9][10]

그러므로 패킷이 p의 확률로 손실되고 1-p의 확률로 손실되지 않는다고 가정하는 uniform model은 실제 모델과는 차이가 난다. 그래서 평균 패킷 손실

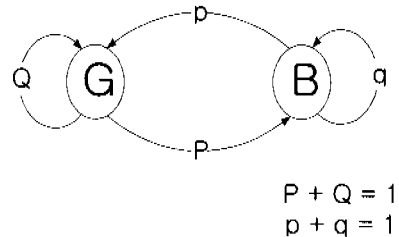


그림 7. Gilbert 모델의 상태 전이도

확률은 uniform 분포와 같으나 연속되는 패킷손실의 길이가 uniform 분포와 다른 경우는 Gilbert model을 적용할 수 있다.<sup>[11]</sup> Gilbert model의 상태 (a) 도는 <그림 7>과 같다.

Gilbert 모델에서 B 상태는 패킷이 손실된 상태를 나타내며, 다른 G 상태는 패킷이 손실되지 않고 목적지까지 도착한 상태를 나타낸다. 현재 패킷이 손 (b) 있었을 경우 다음 패킷이 손실되지 않을 경우의 조건부 확률은 p이고, 다음 패킷도 손실될 조건부 확률은 q이다. 반대로 현재 패킷이 손실되지 않았을 경우 다음 패킷도 손실되지 않을 조건부 확률은 Q 이고, 다음 패킷이 손실될 조건부 확률은 P이다.

(c) Gilbert 모델을 적용하였을 경우 평균 손실 확률은 B에 머무르는 평균시간을 B에 머무르는 평균시간과 G에 머무르는 평균시간의 합으로 나누어 식 (1)과 같이 구할 수 있고, n개의 연속적인 패킷 손실이 일어날 확률은 식(2)와 같다.

$$(d) L_{avg} = \frac{P}{P+p} \tag{1}$$

$$L_{n-continuous} = (1-p)p^{n-1} \tag{2}$$

부가 정보를 가지고 있는 경우 패킷 손실은 2개 이상의 군집 오류가 발생하였을 때와 같이 부가 정보를 이용해 복원하지 못하는 경우에만 발생한다고 볼 수 있으므로 부가 정보를 이용한 손실 패킷 복구후의 손실확률은 식 (3)과 같이 구할 수 있다.

$$L_{recom} = \frac{P(1-p)}{P+p} \tag{3}$$

Gilbert 모델을 적용할 경우 연속 패킷 손실 확률은 <그림 8>과 같다.

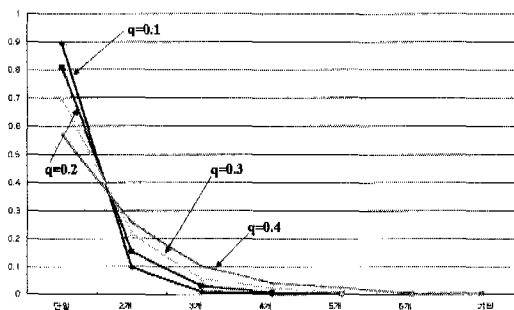


그림 8. 연속 패킷 손실 확률

#### 4.2 실험 및 결과고찰

본 논문에서 제안한 방법의 성능을 평가하기 위해 객관적 음질 평가 방법인 SNR과 주관적 음질 평가를 위해 MOS 테스트를 실시하였으며, 비교되는 부호화기들은 8 kbps G.729(CS-ACELP) 및 AMR 부호화기의 12.2 kbit/s 모드를 포함한다. 평균 패킷 손실률은 0%, 1%, 3%, 5%, 10% 그리고 20%를 사용하였다. 패킷의 크기는 부호화기 자체의 프레임 크기로 하였고, q값은 0.1을 사용하였다. CS-ACELP와 AMR의 12.2 kbit/s 모드에서의 손실 패킷에 대한 처리 방법은 부호화기 자체에 내장된 오류 은닉 방법을 사용하였다.

객관적 성능 평가를 위해 SNR(Signal to Noise Ratio)을 사용하였다. SNR은 식 (4)와 같이 계산되어진다.

$$SNR(dB) = 10 \log \left[ \frac{\sum_{n=1}^N s^2(n)}{\sum_{n=1}^N (s(n) - \hat{s}(n))^2} \right] \tag{4}$$

여기서, s(n)은 원 음성 신호이고  $\hat{s}(n)$ 은 복원 음성 신호이다.

<그림 9>에서 점선으로 표시된 부분이 연속으로 패킷이 손실된 부분인데 제안된 방법이 부호화기 자체에 내장된 오류 은닉방법보다 더 많은 음성을 복구해 냄을 알 수 있다.

평균 패킷 손실에 따른 SNR 값은 <표 2>에 나타나 있다. 실험결과 제안된 방법이 CS-ACELP와 AMR - 12.2 kbit/s 모드보다 우수한 성능을 보였는데, 20%의 평균 패킷 손실에서 제안된 방법이 CS-ACELP 보다는 3.44 dB, AMR - 12.2 kbit/s 모드 보다는 5.71 dB 높다.

표 2. 평균 손실률별 SNR 결과

	CS-ACELP [dB]	AMR-12.2 kbit/s [dB]	제안된 방법 (7.4+4.75) [dB]
0%	15.27	18.92	14.96
1%	12.99	12.77	14.11
3%	9.51	7.55	12.39
5%	8.20	6.53	11.53
10%	6.05	3.49	9.08
20%	3.22	0.95	6.66

주관적 성능 평가는 남녀 각각 6분장으로 총 음성 길이는 1413개의 프레임(1프레임 = 20ms)을 사용했으며, 비전문 청취자 10명을 대상으로 MOS

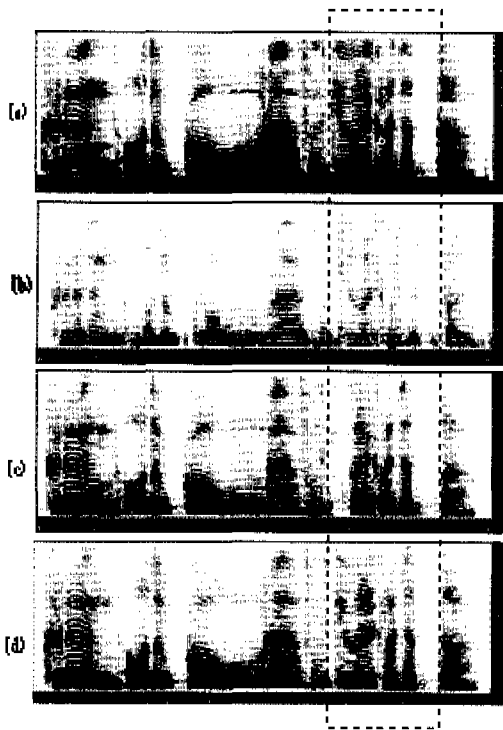


그림 9. 20% 평균 패킷 손실 시 음성 부호화기의 출력 음성의 스펙트로그램  
 (a) 원본 신호  
 (b) CS-ACELP 부호화기  
 (c) AMR - 12.2 kbit/s 부호화기  
 (d) 제안된 방법(7.4kbit/s+4.75kbit/s)

(Mean Opinion Score) 시험이 이루어졌다. <표 3>의 MOS 결과 값도 역시 제안된 방법이 비교 대상 부호화기보다 더 높았다. 20%의 패킷 손실에서 제안된 방법이 CS-ACELP보다는 1.03 높고 AMR-12.2 kbit/s 모드보다는 1.6 높았다.

표 3. MOS 실험 결과

분류 \ 손실률	0%	1%	3%	5%	10%	20%
CS-ACELP	4.07	3.98	3.61	3.45	3.07	2.18
AMR-12.2kbit/s	4.14	3.93	3.38	2.98	2.59	1.88
제안된 방법 (7.4k+4.75k)	4.03	4.00	3.83	3.79	3.51	3.15
원 음	4.3					

### V. 결론

본 논문에서는 AMR 부호화기를 이용하여 인터넷을 통한 음성 전송에 사용하기 위한 방법을 제안

하였다. 인터넷을 기반으로 한 음성 데이터 전송에서 고려해야할 단일 손실 패킷에 대한 복구 방법으로 순방향 오류 정정 방법(FEC)을 사용하였는데, AMR 부호화기의 특징이 다중 전송률을 이용하여 주된 오디오 정보로는 AMR 부호화기의 7.40 kbit/s 모드를 사용하고, 부가 오디오 정보로 AMR 부호화기의 저 전송률인 4.75 kbit/s 모드를 사용하였다. 2 개 이상의 연속 패킷 손실이 발생할 경우, 순방향 오류정정(FEC) 방법으로는 한 개의 패킷만을 복구할 수 있으므로 복구 불가능한 패킷에 대해서는 과거에 올바르게 수신한 패킷과 부가 정보로 복구한 패킷에 보간법을 적용하여 복구 불가능한 패킷을 재생산해 내는 방법으로 오류를 은닉했다. 또한 패킷 손실로 유발된 상태의 비동기로 발생하는 상태 오류를 줄이기 위해 복구 불가능할 경우 부호화기에서 과거 여기 신호를 '0'으로 초기화하는 방법으로 상태 오류를 줄였다. 인터넷에서의 패킷 전송 모델은 Gilbert Model을 사용하였다. 평균 패킷 손실률을 바꿔 가면서 실험한 결과 SNR값 및 MOS 테스트 결과 더 좋은 성능을 보였으며, 20%의 패킷 손실에서 MOS값은 3.2 SNR값은 6.7 dB로 통신 가능한 음질 이상을 보였다. 제안된 방법은 IMT-2000 시스템과 인터넷망과의 음성 통신을 위해 AMR 코더의 전송 손실 보완 방법으로 사용 가능하다.

### 참고 문헌

- [1] H. Schulzrinne, S.Casner, R. Frederick, V. Jacobson, "RTP : A Transport Protocol for Real-Time Applications", RFC-1889, January 1996
- [2] J.-C. Bolot, et al., "Adaptive FEC-based error control for Internet telephony", *Proc. IEEE Infocom '99*, March 1999.
- [3] V. Hardman et al., "Reliable audio for use over the Internet", *Pro. INET 95, Honolulu, HI*, pp. 171-178, June 1995.
- [4] J.-C. Bolot, A. Vega-Garcia, "The case for FEC based error control for packet audio in the Internet", *To appear in ACM Multimedia Systems*.
- [5] C. Perkins, O. Hodson, V.Hardman, "A Survey of Packet Loss Recovery Techniques for Streaming Audio", *IEEE Network*,

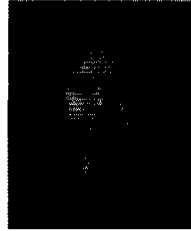


September/October 1998

- [6] M. Handley, "An examination of Mbone performance", *USC/ISI Research Report ISI/RR-97-450*, April 1997.
- [7] M. Yajnik, J. Kurose, D. Towsley, "Packet loss correlation in the Mbone multicast network", *IEEE Global Internet Conference*, November 1996
- [8] H. Schulzrinne, J. Kurose, D. Towsley, "Loss correlation for queues with bursty input streams", *Proc. IEEE ICC '92*, 1992.
- [9] D. Veitch, "Understanding end-to-end Internet traffic dynamics", *IEEE Globecom 98*, November 1998
- [10] M. Yajnik, S. B. Moon, J. Kurose, D. Towsley, "Measurement and modeling of the temporal dependence in packet loss", *IEEE Infocom '99*, March 1999
- [11] Dan Rubenstein, Jim Kurose, Don Towsley, "Real-Time Reliable Multicast Using Proactive Forward Error Correction", *UMASS CMPSCI Technical Report:98-19*, 1998.
- [12] M. Podolsky, C. Romer, S. McConne, "Simulation of FEC-based Error Control for Packet Audio on the Internet", *Proc. IEEE INFOCOM '98*, April 1998
- [13] C. Montminy, T. Aboulnasr, "Improving the performance of ITU-T G.729A for VoIP," *ICME 2000*, August 2000

김 은 중(Eunjoong Kim)

학생회원



2000년 2월: 충북대학교  
전파공학과 졸업  
2000년 3월~현재: 충북대학교  
전파공학과 석사과정  
<주관심 분야> 음성 오디오  
부호화, VoIP

이 인 성(Insung Lee)

정회원



1983년 2월: 연세대학교  
전자공학과 학사  
1985년 2월: 연세대학교  
전자공학과 석사  
1992년 12월: Texas A&M  
University 전기공학과  
박사  
1986년 5월~1987년 7월: 한국통신 연구개발단  
전임연구원  
1993년 2월~1995년 9월: 한국전자통신연구원 이동  
통신기술연구단 선임연구원  
1995년~현재: 충북대학교 전기전자 및 컴퓨터  
공학부 부교수  
<주관심 분야> 음성 및 영상신호압축, 이동통신, 적  
응필터