

논문 2005-42TC-12-17

CELP 부호화기를 위한 양방향 패킷 손실 은닉 알고리즘

(BS-PLC(Both Side-Packet Loss Concealment) for CELP Coder)

이 인 성*, 황 정 준**, 정 규 혁**

(In-Sung Lee, Jeong-Joon Hwang, and Gyu-Hyeok Jeong)

요 약

VoIP에서 패킷의 손실은 음성의 품질에 영향을 주는 가장 중요한 요인이다. 따라서 수신된 정보로부터 손실된 패킷을 복구하는 것은 중요하다. 따라서 본 논문은 VoIP에서 가장 많이 사용되는 CELP 부호화기를 위한 수신측 기반의 손실 패킷 복구 방법을 제안한다. 제안하는 WSOLA(Waveform Shift OverLab Add)기반의 BS-PLC (Both Side Packet Loss Concealment) 방법은 패킷 손실이 발생하였을 경우 미래 패킷을 이용할 수 있는 경우와 그렇지 않을 경우로 나누고, LP(Linear prediction) 파라미터와 여기 신호를 복구한다. 미래 패킷을 이용할 수 없는 경우에는 과거에 전송된 정상 패킷만을 가지고 복원을 하며, 미래 패킷을 이용할 수 있을 경우에는 과거의 정상 패킷과 미래의 정상 패킷을 동시에 이용하여 손실된 패킷을 복구한다. 연속 패킷 손실 환경은 Gilbert 모델로 설정하였고, 제안한 알고리즘을 VoIP에서 가장 많이 사용되는 CELP 음성부호화기인 G.729에 적용하여 성능을 비교한다. 성능 비교를 위해 손실율을 변화시키면서 SNR(Signal to Noise)와 MOS(Mean Opinion Score)측정하였고, 제안한 방법을 G.729의 패킷 손실 은닉 방법과 비교하였다. 실험 결과, 평균 손실률이 20%에서도 SNR은 2dB, MOS값은 0.3정도도의 음질 개선을 보였다.

Abstract

Lost packet robustness is an most important quality measure for voice over IP networks(VoIP). Recovery of the lost packet from the received information is crucial to realize this robustness. So, this paper proposes the lost packet recovery method from the received information for real-time communication for CELP coder. The proposed BS-PLC (Both Side Packet Loss Concealment) based WSOLA(Waveform Shift OverLab Add) allow the lost packet to be recovered from both the "previous" and "next" good packet as the LP parameter and the excitation signal are respectively recovered. The burst of packet loss is modeled by Gilbert model. The proposed scheme is applied to G.729 most used in VoIP and is evaluated through the SNR(signal to noise) and the MOS(Mean Opinion Score) test. As a simulation result, The proposed scheme provide 0.3 higher in Mean Opinion Score and 2 dB higher in terms of SNR than an error concealment procedure in the decoder of G.729 at 20% average packet loss rate.

Keywords : VoIP, 에러은닉, G.729, CELP, 음성부호화기

I. 서 론

음성 정보를 인터넷망을 통해 전송하는 기술인 VoIP에서는 실시간으로 데이터를 전송하기 위해 TCP/IP와는 달리 UDP/RTP/IP를 사용한다. UDP는 최소한의 데이터를 전송하는 프로토콜로써 TCP에서 제공하는 시간만료, 재조립, 재전송, 에러검출 등의 기능을 포함하

고 있지 않다. 따라서 UDP/RTP/IP를 사용하는 VoIP에서는, 데이터 전송 중 종단간 지연(End-to-End Delay), 지터(Jitter), 손실(Packet Loss), 비순차 패킷(Out-of-Order Packet) 등이 발생하여 실시간 음성 통신의 품질을 보장받기 어렵다^[1]. 위와 같은 성능 파라미터들 중 종단간 지연과 비순차 패킷은 수신측에서 버퍼링(buffering)을 하고, 재생 시간을 지연시킴으로써 상당수 해결할 수 있지만, 복합적인 영향으로 결국에는 패킷 손실로 이어져 복호화 음성이 끊어지거나 일그러지고 심지어는 통신 장애가 생기기까지 한다. 따라서 패킷손실 특성과 패킷손실복구 방법에 관한 연구의 필요성이 증가하게 되었다^{[2][3][4]}

* 정희원, ** 학생회원, 충북대학교 전자공학과
(Dept. of Radio Science & Engineering Chungbuk National University)

※ 본 연구는 2005년도 충북대학교 학술연구 지원사업의 연구비 지원으로 수행되었습니다.

접수일자: 2005년9월8일, 수정완료일: 2005년12월5일

CELP기반의 표준 음성부호화기의 부록으로 제공되는 패킷손실복구방법^{[5][6][7]}은 패킷 손실에 대해서 피치 주기나 LP계수를 반복하여 복구 및 은닉을 한다. 하지만 패킷손실이 연속적으로 발생할 경우 같은 음성신호를 반복함으로써 오히려 부자연스러운 인공음을 복원하게 된다. 이를 보완하기 위해 이득 감쇄를 적용했지만 몇 개 이상의 연속 패킷 손실이후의 정상 프레임과의 연결성이 부자연스러워지게 되고, 정상 프레임마저 신호가 왜곡되는 문제점이 있다. 또한, 음소가 변화하는 구간에서 단일 패킷 손실이 발생할 경우에는 과거 패킷만을 이용하여 보상하는 패킷손실복구 알고리즘은 음절 저하를 초래하게 된다.

본 논문은 이 같은 문제에 초점을 맞추어 손실 프레임의 과거 패킷과 미래 패킷을 이용하여 LP 파라미터와 여기신호를 각각 보간하는 BS-PLC방법을 제안한다. 제안하는 손실 패킷 복원 및 에러 은닉 방법은 전송률 부가 정보를 삽입하는 FEC(Forward Error Correction) 방법^[8]과는 달리 대역폭의 증가를 요구하지 않는다. 또한 손실된 패킷 복원 시 부호화기의 비트스트림 정보를 이용하므로 많은 연산과정이 필요하지 않는 장점이 있다. 단점이라 할 수 있는 추가적인 지연은 미래 패킷을 이용할 수 없는 경우에 과거 패킷만을 이용함으로써 최소화시킨다.

본 논문은 총 V장으로 구성된다. II장에서는 VoIP에서 주로 사용되고 있는 G.729 음성부호화 알고리즘과 부호화기 내장된 패킷 손실 복구 방법에 대해 설명하고, III장에서는 새로운 패킷 손실 복구 방법을 제안한다. IV장에서는 제안하는 방법이 적용된 G.729 음성 부호화기의 성능을 손실률의 변화에 따라 SNR(Signal to Noise)과 MOS(Mean Opinion Score)측정함으로써 G.729에 내장된 패킷손실복구 방법과 비교한다. 이때의 패킷손실환경은 길버트 모델로 설정한다. 마지막으로 V장에서는 결론을 맺는다.

표 1. G.729 음성부호화기의 비트 할당
Table 1. Bit allocation of the G.729.

LSP	L0, L1, L2, L3			18
적용코드북지연	p1, P2	8	5	13
피치 지연 parity	P0	1		1
고정 코드북 인덱스	C1, C2	13	13	26
고정 코드북 부호	S1, S2	4	4	8
코드북 이득 1	GA1, GA2	3	3	6
코드북 이득 2	GB1, GB2	4	4	8
Total				80

II. G.729 음성부호화기^{[5][6]}

G.729 음성부호화기는 ACELP 기반의 부호화기 로써 8 kbit/s의 전송률을 가진다. 이 부호화기는 0Hz~3400Hz 대역의 음성신호를 8kHz로 표본화된 16 bit linear PCM데이터를 입력 신호로 가진다.

G.729 음성부호화기는 음성 프레임의 길이가 5ms인 2개의 부 프레임을 갖는 10ms 음성 프레임 마다 분석된다. 추출된 파라미터들은 프레임 당 80 비트가 할당되어 부호화된다.

부호화 과정에서 사용되는 모델 파라미터와 비트할당은 표 1에 나타나있다. 그림 1은 G.729의 전체적인 부호화 과정을 보여준다. 입력되는 음성신호는 고역필터를 통과한 후 자기상관법에 의해 선형예측계수(LPC)가 추출된다. LPC 계수의 계산 및 전송은 10ms 프레임 마다 한번 수행된다. 음성신호의 스펙트럼 포락선 정보를 나타내는 LPC 계수는 전송을 위하여, LSP(Line spectrum Pairs) 파라미터 값으로 변환된다. LSP 파라미터의 전송에는 18 비트가 할당되며, 파라미터 값들은 2단 예측 벡터 양자화된다.

적용 코드북과 고정 코드북 파라미터 값들은 5ms 부 프레임마다 검색되어 전송된다. 피치 지연의 검색 시 계산량을 줄이기 위하여 모든 피치지연에 대하여 폐구간 검색을 하지 않고, 개구간 방법에 의하여 예측된 피치 지연의 주변 값들만을 검색하여 정확한 피치 지연값을 찾아낸다. 검색되어지는 피치지연은 최대 1/3 샘플을 갖는다. 피치지연 값은 첫 번째 부 프레임에서는 8 비트로 부호화되고, 두 번째 부 프레임에서 앞 부 프레임 지연 값의 차이를 5 비트를 사용하여 부호화한다.

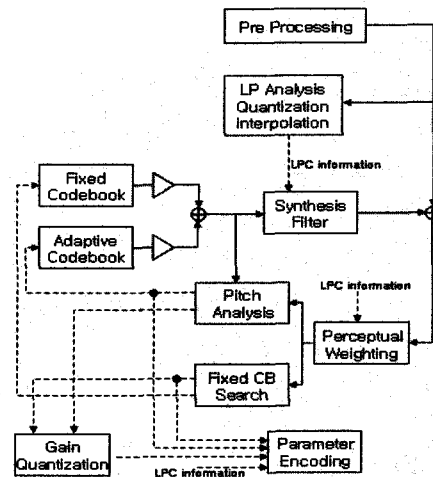


그림 1. G.729 인코더 블록다이어그램
Fig. 1. Block diagram of the G.729.

최적의 피치 지연 값을 찾아낸 후, 음성의 여기신호 추출을 위한 고정 코드북 검색이 수행된다. 고정 코드북 파라미터는 5ms 부 프레임 마다 분석/합성의 방법으로 검색된다. 고정 코드북 검색을 위한 목적 신호는 선형예측 잔여신호에서 적응 코드북의 기여분을 제거함으로써 얻어진다. 고정 코드북 여기신호를 발생은 대수적 코드북을 사용한다. 고정 코드북의 비트할당은 펄스 위치에 13 비트, 펄스 부호에 4 비트를 할당하며, 전체적으로 17 비트를 할당한다. 사용되는 코드북 구조는 각 부 프레임마다 4개의 펄스만이 지정된 위치에 +1, -1의 값을 갖게 되므로 고정 코드북 검색 시 계산량을 줄일 수 있다. 적응 코드북과 고정 코드북의 이득 양자화에는 Conjugate 구조를 갖는 2개의 코드북을 사용한다. 각각의 코드북은 8개와 16개의 요소 값을 가지며 구해진 이득 값과 미리 선택된 4개와 8개의 요소 값들에 대해서만 검색이 이루어지므로 계산량을 줄일 수 있다. 또한 프레임간의 상관성을 이용한 4차 MA(Moving average) 예측기를 사용한다. 부호화의 최종 단계에서는 다음 프레임의 목적 신호를 구하기 위해 합성 및 가중 필터의 메모리 갱신을 수행한다.

프레임 손실을 검출하는 기법은 권고안에 정의되어 있지 않고, 응용에 따라 달라진다. 먼저 패킷이 손실되면 이전 정상 프레임에서의 LP 파라미터 값들을 그대로 반복하여 LP 합성 필터를 만든다. 다음 프레임과의 연속성을 유지하기 위해 적응 코드북 이득과 고정 코드북 이득은 0.9와 0.98의 비율로 감쇄된다. 이득 예측기의 메모리도 감쇄시킨다. 대체 여기신호는 유성음과 무성음을 구별하여 달리 생성된다. 이득 예측기의 결과를 이용하여 이득의 크기가 3dB 이상일 경우 유성음으로 판별하고, 이때의 여기신호는 적응 코드북 이득만을 이용하여 여기신호를 생성한다. 반면 이득의 크기가 3dB 미만일 경우 무성음으로 판별 후 고정 코드북의 이득만을 이용하여 고정 코드북에서 랜덤하게 발생시킨다.

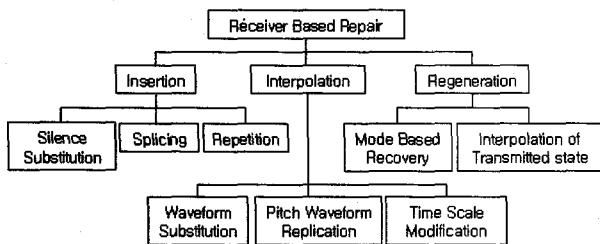


그림 2. 수신단 기반의 패킷손실복구방법의 분류
 Fig. 2. Classification of the packet loss recovery method for seceiver based repair.

이렇게 만들어진 대체 여기 신호는 LP 파라미터가 반복 되어 만들어진 합성필터를 통과시켜 손실 패킷을 복원 한다.

III. 패킷 손실 복구 방법

1. 인터넷에서의 패킷 손실 복구 방법

패킷 복구 및 오류 은닉 기술은 주제에 따라 송신단 기반 복구 방법과 수신단 기반 오류 은닉 방법으로 나눌 수 있다. VoIP에 많이 사용되고 있는 기존의 부호화 기술(G.711^[9], G.723.1^[7], G.729^[6])들의 오류 은닉 방법은 송신단으로부터 어떤 도움도 받지 않는 수신단 기반의 오류은닉을 사용한다. 이것은 송신자 기반의 오류은닉 방법이 송신단 뿐만 아니라 수신단의 장비를 교체하거나 수정할 필요가 없기 때문이다.

그림 2와 같이 수신단 기반의 패킷 손실의 복구 기술은 크게 삽입 기반의 오류은닉 방법, 보간법 기반의 오류은닉 방법 그리고 재생 기반의 오류은닉 방법으로 나뉘어진다.

삽입 기반의 오류 은닉 방법은 손실 프레임을 묵음 혹은 이전 프레임으로 대신하거나 손실 프레임을 빼버리는 방법으로 계산 복잡도가 매우 낮지만 성능 또한 매우 낮다. 보간법 기반의 오류 은닉 방법은 음성신호의 특성 변화를 고려한 방법으로써 삽입기반의 은닉 방법보다 계산 복잡도는 조금 더 높으며, 성능이 보다 뛰어나다. 재생산 기반의 오류 은닉 방법은 부호화기에 의존적이며, 많은 수의 음성 정보가 복구에 사용되므로 성능이 매우 좋지만 계산 복잡도가 매우 높다^{[10][11]}.

따라서, 실시간 음성 통신 서비스를 구현하기 위해 추가적으로 전송 지연을 발생 시키지 않는 수신단 기반의 에러 은닉 방법이 적합하지만 손실된 패킷에 대한 어떠한 정보도 가지지 못하므로 손실률이 낮고(15% 이하) 단구간(4~40ms)의 손실에서 적용이 가능하다^[12]. 과거의 정보를 너무 오랫동안 사용하여 손실을 보상한다면, 인공음에 의해 복원된 음성은 오히려 더 부자연스러워진다. 이 같은 이유로 대부분의 에러 은닉 알고리즘은 신호의 감쇄를 적용하여 60ms 이후에는 묵음으로 대체하는 방법을 사용한다^[9]. 하지만 복호화기에서 음성을 재생하는 동안에도 다음 음성 프레임들은 지터 버퍼에 저장되고 있으므로, 버스트(burst)에러 등에 의해 연속손실이 크게 발생하였을 때 손실구간 뒷부분은 손실 이후 프레임 정보로 보상하는 것이 바람직하다.

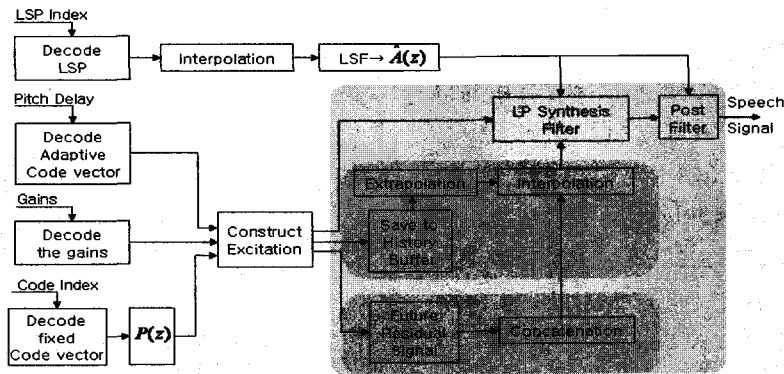


그림 3. 제안하는 BS-PLC 알고리즘이 적용된 G.729 음성복호화기 블록도
 Fig. 3. Block diagram of G.729 decoder applying the proposed BS-PLC.

2. 제안하는 양방향 패킷 손실 은닉 알고리즘

제안한 손실 패킷 복구 및 에러 은닉 방법을 G.729에 적용한 구조는 그림 3과 같다. 현재 패킷이 오류로 검출되거나 손실되면 오류 미래 패킷을 이용할 수 있는지에 따라 그림 4와 같은 방법으로 손실 패킷을 복원한다. 그림 4의 점선 네모블록은 미래 패킷을 사용할 수 있는 경우를 나타낸다.

가. 미래 패킷을 이용할 수 없는 경우

미래 패킷을 이용 할 수 없는 경우에는 과거 버퍼의 데이터만을 가지고 손실 패킷을 은닉한다. 이때의 패킷 손실 복원과정은 그림 4에서 점선 네모블록을 제외한 나머지 과정을 거친다.

(1) LP(Linear Prediction) 파라미터 보간

LP(Linear Prediction) 합성필터의 대역폭을 확장한다는 것은 필터의 피크를 약간씩 줄이고 대역폭을 조금 늘리는 것을 의미한다.

$$a'_i = (0.98)^i a_i, \quad i = 1, \dots, L \tag{1}$$

식 (1)을 적용함으로써 이전 프레임으로부터 구해진 LP 합성필터는 조금 더 부드러운 형태가 된다^[13]. t 는 연속 패킷 손실된 수이고 L 의 값은 10이며 선형예측계수의 수이다.

이로 인해 천천히 변화하는 음성신호의 스펙트럼 특성을 반영할 수 있다. 패킷 손실이 발생하는 때 프레임마다 LP 합성 필터의 대역폭을 확장하여 필터의 스펙트럼을 평활화 시키고, 합성음성을 보다 부드러운 음성 변화구간을 얻을 수 있다.

(2) 여기신호 복원

손실된 패킷의 피치주기는 수신된 비트스트림에 이 미 G.729 음성부호화기에서 구한 피치정보를 이용하여 검색 구간을 20샘플로 결정하고 교차상관도를 구한다. 과거 신호의 마지막 부 프레임의 피치주기를 적용하여 그림 5와 같이 OLA(OverLap Add)한다. 이는 과거에 수신된 피치 정보로 검색구간을 정함으로써 시스템의 복잡도를 줄인다. 뿐만 아니라 G.711 PLC에서 2:1로 축약하여 검색할 때 생기는 피치 더블링에 의한 음질저하가 발생할 수 있는 가능성을 줄일 수 있다^[9].

$$\phi_{AB} = \sum_{j=1}^{160} A_j B_{ij} \tag{2}$$

손실 패킷의 여기신호의 생성은 위에서 찾은 피치를 기반으로 피치주기를 반복하여 사용한다. 처음 10ms는 피치주기를 그대로 반복하여 사용한다. 이후 20ms에서 두 배로 증가된 피치 주기를 사용하고, 30ms에는 피치 주기가 한번 더 더해져 세배가 된다. 그 이후 손실에 대해서는 피치 버퍼에서 추가적인 수정없이 만든다. 이는

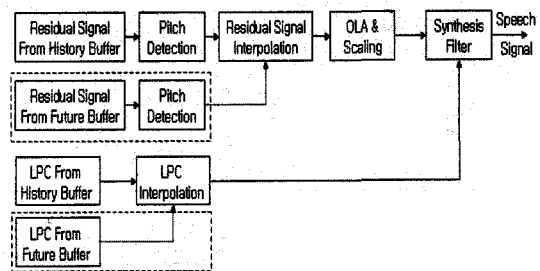


그림 4. 제안한 손실패킷 복구 및 에러 은닉 구조
 Fig. 4. The proposed scheme for recovery of lost packet and error concealment.

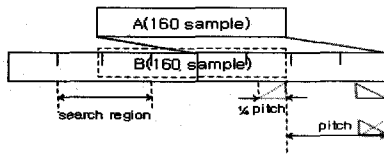


그림 5. 교차상관도를 이용하여 피치 결정
Fig. 5. Pitch detection using cross-correlation.

손실 구간 처리 중에 신호를 합성하기 위해 사용된 피치 주기의 수가 증가할수록 부자연스러운 소리가 감소된다는 것을 실험적으로 알 수 있기 때문이다.

OLA는 과거 버퍼와 피치를 이용하여 복원한 패킷 사이에 연속성을 유지하기 위하여 수행한다. 과거 버퍼의 출력되어지지 않은 음성신호의 마지막 20 샘플(2.5ms)

과 복원된 음성신호 앞의 20 샘플(2.5 ms)은 삼각 윈도우를 이용하여 가중치가 적용되어 더해진다. 더해진 신호는 과거 버퍼의 마지막 20 샘플(2.5 ms)을 대체한다. 과거 버퍼는 복원된 신호를 저장하기 위해 80 샘플(1 frame) 만큼 left-shift 되어 지고 마지막 80 샘플을 복원된 신호로 대체한다.

손실 발생시 스케일 감쇄는 [9]에서 서술되어 있는 것과 같다. 스케일은 초기에 1.0으로 설정된다. 만일 손실이 발생이 되면 첫 패킷에서부터 20ms까지 10ms 당 0.054씩 감소가 된다. 그리고 20ms 이상 손실부터는 10ms당 0.222씩 감소가 되어 최종적으로 60ms 연속 손실부터는 0이 된다.

나. 미래 패킷을 이용할 수 있는 경우

만일 미래 프레임을 사용할 수 있다면 제안된 패킷손실은닉 알고리즘은 그림 6과 같이 손실 패킷을 복원하기 위해 미래 패킷 정보와 과거 패킷 정보를 동시에 이용한다.

(1) LP(Linear Prediction) 파라미터 보간

만일, n번째 패킷이 손실되고 미래 패킷인 n+1번째

표 3. LP(Linear Prediction) 파라미터 보간
Table 3. Interpolation of LP(Linear Prediction) parameters.

	n 번째 프레임(i = 1, ..., L)	n-1 번째 프레임(i = 1, ..., L)
부 프레임 1	$a_i(n)^{(1)} = 0.75a_i(n-1) + 0.25a_i(n+1)$	$a_i(n+1)^{(1)} = 0.25a_i(n-1) + 0.75a_i(n+1)$
부 프레임 2	$a_i(n)^{(2)} = 0.5a_i(n-1) + 0.5a_i(n+1)$	$a_i(n+1)^{(2)} = a_i(n+1)$

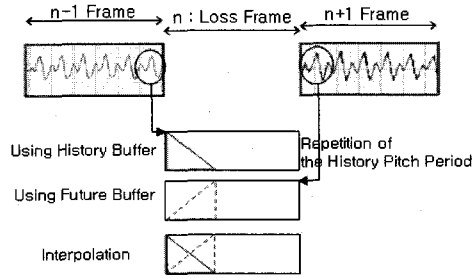


그림 6. 과거 신호와 미래 신호 보간
Fig. 6. Interpolation between the signal from history buffer and the signal from future buffer.

패킷을 이용 할 수 있다면, 손실된 n+1번째 패킷의 LP 파라미터 값들은 표 2와 같이 각각 보간 된다.

(2) 여기신호 복원

과거 버퍼로부터 피치를 찾는 과정은 과거 버퍼만을 이용한 방법과 동일하게 수행된다. 미래 버퍼로부터 피치를 얻기 위해서 미래 프레임의 첫 번째 부 프레임의 피치주기를 적용한다.

피치는 미래 신호의 앞에서부터 선택한다. 선택된 피치를 미래 신호 앞에 연결을 하게 되므로 신호의 연속성을 주기 위해 미래 신호의 한 피치주기 이후 신호와 미래 신호의 시작 부분의 신호를 미래 피치의 1/4만큼 OLA를 적용하여 사용한다. 다음과 같이 결정된 피치를 반복 사용하여 마지막 손실 패킷을 보간하기 위한 한 프레임을 만든다.

미래 신호를 사용하지 않고, 과거 신호만을 이용하여 손실된 패킷을 보간 하기 위해 또 다른 한 프레임의 음성신호를 생성한다. 미래 신호를 이용하여 생성한 여기신호와 과거 버퍼를 이용하여 생성한 여기신호를 그림 6과 같이 OLA를 적용하여 사용한다. 이때 OLA의 길이는 식(3)과 같이 손실 패킷 길이에 따라 적용된다.

$$\begin{aligned} \text{overlaplength} &= T_p^{\text{future}}/4 + E_{\text{cnt}} \times 32 (4\text{ms}) \\ (\text{overlaplength} \leq 80\text{samples}) \end{aligned} \quad (3)$$

여기서 T_p^{future} 는 미래 신호의 피치주기이고, E_{ent} 는 손실된 프레임의 개수를 나타낸다.

OLA를 적용한 후 이후 음성신호는 미래버퍼의 음성신호를 채워 넣는다. 손실 이후 첫 번째 정상적으로 전송된 패킷에서 복원된 신호와 정상적인 신호 사이의 부드러운 변이가 필요하다. 만일 스케일 변수가 1.0보다 작으면, 출력신호는 스케일 업 된다. 이때 증가율은 10ms 당 0.498씩 증가한다. 이 과정은 출력 프레임의 마지막 신호 혹은 스케일이 1.0이 될 때 까지 계속된다.

IV. 시뮬레이션 및 성능 평가

1. 실험 환경

인터넷 환경에서의 패킷 손실과정의 상관구조를 길버트 모델과 같은 저 차의 Markov 체인으로 나타낼 수 있다는 연구결과가 나왔다^[14]. 즉 길버트 모델은 그림 7과 같이 인터넷상의 패킷 손실과정을 일종의 on-off Markov process로 모델링 한 경우이다. 손실은 B상태에서만 일어나며 이전 프레임이 G 상태일 때 손실이 발생할 확률이 P이고, 이전 프레임이 B상태일 때 손실이 발생할 확률은 q이다. 길버트 모델을 적용하였을 경우 평균 손실 확률은 B 상태에 머무르는 평균시간을 B 상태에 머무르는 평균시간과 G상태에 머무르는 평균시간의 합으로 나누어 식 (4)와 같이 계산할 수 있고, 패킷손실의 평균길이는 식 (5)와 같이 구할 수 있다.

$$P(Loss) = \frac{P}{P+q} \tag{4}$$

$$E(Loss\ Burst) = \frac{1}{p} \tag{5}$$

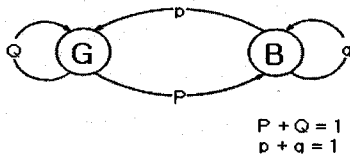


그림 7. 길버트 모델의 상태 천이도
 Fig. 7. State diagram of the Gilbert model.

2. 실험 및 결과 고찰

논문에서 제안한 방법의 성능을 평가하기 위해 객관적 음질 평가 방법인 SNR과 주관적 음질 평가를 위해 MOS 테스트를 실시하였고, 제안한 패킷손실복구 방법을 G.729 부호화기의 오류은닉 방법과 비교한다.

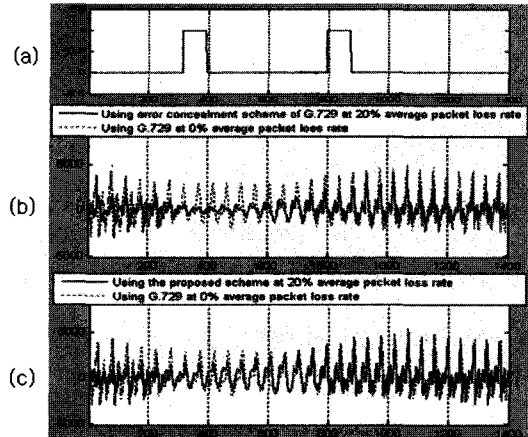


그림 8. 단일 패킷 손실에서 복원된 음성신호 a) 손실구간 b) G.729 자체 에러 은닉 사용하여 복원 c) BS-PLC 알고리즘을 적용하여 복원

Fig. 8. Recovered signal at single packet loss. a) Frame of packet loss b) Using error concealment in G.729 decoder c) Using the proposed BS-PLC.

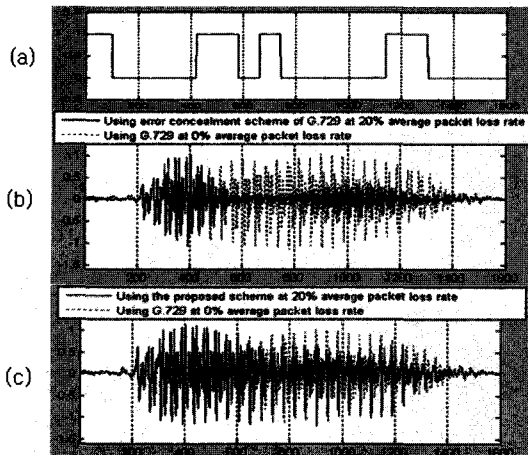


그림 9. 연속 패킷 손실에서 복원된 음성신호 a) 손실구간 b) G.729 자체 에러 은닉 사용하여 복원 c) BS-PLC 알고리즘을 적용하여 복원

Fig. 9. Recovered signal at bursted packet loss. a) Frame of packet loss b) Using error concealment in G.729 decoder c) Using the proposed BS-PLC.

그림 8과 9는 패킷 손실이 일어날 경우, 제안한 방법으로 패킷을 복원한 결과와 G.729의 자체 오류 은닉 방법으로 복원한 결과를 비교한 것이다. 그림 8과 9의 (a)는 손실된 구간을 출력한 것이고, (b)와 (c)에서 점선 파형은 패킷의 손실이 전혀 없을 경우를, 실선 파형은 제안한 방법 또는 부호화기 자체에 내장된 오류 은닉 방법으로 복원한 음성신호를 나타낸다.

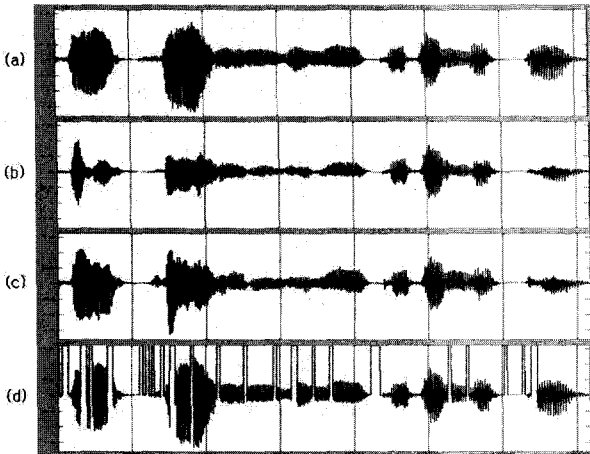


그림 10. 평균 손실률 20%일때 복원된 출력 신호 a) 평균 손실률 0%일때 음성신호 b) G.729 자체 에러 은닉 사용하여 복원 c) BS-PLC 알고리즘을 적용하여 복원 d) 손실 구간

Fig. 10. Output signal at 20% average packet loss rate. a) Speech signal at 0% average packet loss rate b) Using error concealment in G.729 decoder c) Using the proposed BS-PLC d) Frame of packet loss.

그림 8은 단일 패킷 손실이 일어날 경우의 복원정도를 보여준다. 그림 8의 (b)에서 400에서 800 샘플 사이의 신호를 보면 단순히 코드북 이득과 이득 예측기의 메모리를 감쇄하는 것은 손실 구간 다음에 오는 정상 구간에서도 파형의 이득이 작아지게 되어 신호가 왜곡되는 문제점이 있다. 연속 패킷 손실이 일어난 후의 정상 패킷 구간에서는 신호왜곡이 더 심해서 음질 저하에 심각한 영향을 미치게 된다. 그림 9의 (b)의 약 600 샘플에서 1100 샘플 사이의 구간에서 쉽게 알 수 있다. 그림 8의 (c)와 그림 9의 (c)를 보면 이러한 신호 왜곡이 제안한 방법을 이용하여 줄일 수 있음을 알 수 있다.

그림 10은 복원된 출력음성을 나타낸다. 그림 10의 (d)는 손실된 구간을 나타내는데 제안된 방법이 부호화기에 내장된 오류 은닉방법보다 0% 패킷 손실률일 때의 출력음성에 보다 가깝게 복구해 냄을 알 수 있다.

성능 평가에 사용된 음성은 일본 NTT에서 제공하는 데이터베이스에서 남녀 각각 5문장을 사용하였고, 총 2670개의 프레임(1프레임 = 10ms)을 사용하였다. 패킷 손실 환경은 길버트 모델로 설정하였고, 연속 패킷 손실률을 20%로 고정시키고, 평균 패킷 손실률을 0%, 3%, 5%, 10%, 15%, 20%, 그리고 30%를 사용하였다.

MOS 측정을 위해서는 "PESQ - ITU-T Recommendation P.862 Version 1.2 - 2 August 2002"를 사용하였고, SNR은 식 (6)과 같이 계산되어진다.

표 3. SNR 값과 MOS 값 측정 결과
Table 3. Result of SNR and MOS tests.

평균 손실률	G.729 내장된 오류은닉 방법		제안된 오류은닉 방법	
	MOS	SNR(dB)	MOS	SNR(dB)
0%	3.687	15.38	3.687	15.38
3%	3.362	13.57	3.302	13.98
5%	3.128	13.04	3.246	13.09
10%	2.702	9.264	2.867	10.65
15%	2.538	8.105	2.769	9.361
20%	2.237	6.009	2.597	7.993

식 (6)에서 $s(n)$ 은 원 음성신호이고, $\hat{s}(n)$ 은 복원 음성 신호이다.

$$SNR(dB) = 10 \log \left[\frac{\sum_{n=1}^N s^2(n)}{\sum_{n=1}^N \{s(n) - \hat{s}(n)\}^2} \right] \quad (6)$$

평균 패킷 손실에 따른 MOS 결과 값과 SNR 값은 표 3에 나타나 있다. 실험 결과 평균 손실률이 증가함에 따라 제안된 방법이 G.729의 내장된 오류 은닉 방법보다 우수한 성능을 보였다. 20%의 평균 패킷 손실에서 제안된 방법이 SNR 값은 2 dB정도의 향상을 보였고, MOS 값은 0.3정도 높다.

V. 결 론

본 논문에서는 수신단에서 손실된 패킷을 복구하고 에러를 은닉하는 방법으로 BS-PLC 알고리즘을 제안한다. 패킷 손실이 일어날 경우, 수신단에서 패킷이 손실된 프레임의 과거 정상 패킷과 미래 정상 패킷을 이용하여 LP 파라미터와 여기신호를 각각 보간하여 손실된 패킷을 복원한다. 미래 값을 이용하여 복원하므로 G.729에 내장된 오류은닉 방법을 적용하였을 때 복원하지 못한 부분도 복원하였으며 손실 구간 후 정상 구간의 신호 왜곡을 줄일 수 있다. 또한 마지막 손실 부분부터 스케일 업 시킬 수 있어 무손실일 때의 출력 신호와 보다 유사하게 복원이 되었다. 이러한 방법은 부가 정보를 삽입하는 FEC 방법과는 달리 대역폭을 증가시키지 않으며 미래 패킷을 이용할 때의 추가적인 지연은 미래 패킷을 이용할 수 없는 경우에 과거 패킷만을 이용함으로써 최소화시킨다. 또한 수신된 비트스트림 정보를 이용함으로써 많은 연산과정이 필요하지 않는 장점이 있다.

실험 결과, 제안된 패킷손실복구 방법은 G.729 부호화기에 내장된 방법보다 평균 손실률 20%의 패킷 손실

에서 SNR값은 2dB, MOS 값은 0.3 정도의 음질 개선을 보였다.

참고 문헌

- [1] A. Kos, B. Klepec and S. Tomazic, "Techniques for performance improvement of VoIP applications," Electrotechnical Conference, MELECON 11th Mediterranean, pp. 250-254, May 2002.
- [2] J. C. Bolot and A. Vega-Garcia, "The Case for FEC-based Error Control for Packet Audio in the Internet," ACM Multimedia Systems, Jan. 1997.
- [3] 황정준, 이인성, "FEC기능을 추가한 AMR-WB 음성 부호화기를 이용한 음성 패킷 전송," 전자공학 회논문지, 제40권, TC편, 제 11호, pp. 63-71, 11. 2003.
- [4] M. Yajnik, J. Kurose, D. Towsley, "Packet loss correlation in the Mbone multicast network.," IEEE Global Internet Conference, GLOBECOM '96. 'Communications, pp. 18-22, Nov. 1996.
- [5] ITU-T Recommendation G.729, "Coding of speech at 8 kbps using conjugate-structure algebraic code excited linear prediction (CS-ACELP)," Jan. 1995.
- [6] ITU-T G.729 Annex B, "A silence compression scheme for G.729 optimizes for terminals conforming to recommendation V. 70," Nov. 1996.
- [7] ITU-T G.723.1 Annex A, "Silence compression scheme," Nov. 1996.
- [8] N. Shacham and P. McKenney, "Packet Recovery in High-speed Networks using Coding and Buffer Management," Ninth Annual Joint Conference of the IEEE Computer and Communication Societies, INFOCOM '90, San Fransisco, CA, pp 124-131, May 1990.
- [9] ANSI Recommendation T1.521a-2000, "Packet Loss Concealment for use with ITU-T Recommendation G.711," Jul. 2000.
- [10] N. S. Jayant and S. W. Christenssen, "Effects of packet losses in waveform coded speech and improvements due to an odd-even sample-interpolation procedure," IEEE Transactions on Communications, vol. 29, no 2, pp. 101-109, Feb. 1981.
- [11] R. M. Warren, "Auditory Perception," Pergamon Press, Inc., 1982.
- [12] 김무중, 권홍철, "인터넷 전화에서 손실 패킷 복원을 위한 동적인 부가정보 전송 기법," 한국음향학회지, 제21권, 제 4호, pp. 346-360, 2002.
- [13] J. Lindblom and P. Hedelin, "Packet loss concealment based on sinusoidal extrapolation," IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '02, vol .1, pp.173-176, 2002.
- [14] D. Rubenstein, J. Kurose and D. Towsley, "Real-time Reliable Multicast Using Proactive Forward Error Correction," UMASS CMPSCI Technical Report : 98-99, 1998.

저자 소개



이인성(정회원)

1983년 2월 연세대학교
전자공학과(공학사)

1985년 2월 연세대학교
전자공학과 (공학석사)

1992년 2월 Texas A&M
University 전기공학과
(공학박사)

1993년 2월~1995년 9월 : 한국전자 통신연구원
이동통신 기술연구단 선임연구원

1995년 10월~현재 충북대학교 전기전자공학부
정교수

<주관심분야 : 음성/영상 신호 압축, 이동통신,
적용필터>



황정준(학생회원)

2002년 2월 충북대학교 전기전자
공학(공학사).

2002년 3월~2004년 2월
충북대학교 전파통신
공학과 (공학석사)

<주관심분야 : 디지털 신호처리,
음성 신호처리, VoIP>



정규혁(학생회원)

2004년 2월 충북대학교 전기전자
공학(공학사)

2004년 3월~현재 충북대학교
전파통신공학과
(공학석사)

<주관심분야 : 음성/오디오 부호
화, 통신신호처리, VoIP>