

CELP Type Vocoder에서 RTP 확장 헤더 데이터를 이용한 연속적인 프레임 손실에 대한 PLC 성능개선

The Performance Improvement of PLC by Using RTP Extension Header Data for Consecutive Frame Loss Condition in CELP Type Vocoder

홍 성 훈*, 배 명 진*

(Seong-Hoon Hong*, Myung-Jin Bae*)

*송실대학교 정보통신공학과

(접수일자: 2009년 10월 20일; 채택일자: 2009년 12월 1일)

패킷네트워크에서 사용하는 음성부호화기는 자체적으로 PLC (Packet Loss Concealment) 알고리즘을 사용하고 있지만 서비스에 적합할 만큼 좋지 않다. 더욱이 연속적인 패킷 손실에 대해서는 많이 취약하다. PLC 알고리즘은 크게 송신단 기반의 알고리즘과 수신단 기반의 알고리즘으로 나뉜다. 송신단 기반의 알고리즘은 추가정보를 전송하기 때문에 음질개선에 큰 성능을 갖는 반면 데이터 전송율이 높아지고 추가 지연이 발생하며 상호간의 호환이 불가능하다. 수신단 기반의 알고리즘은 수신된 데이터에 기반하기 때문에 추가지연이나 정보가 필요 없으나 음질개선에 한계가 있다. 본 논문에서는 RTP 헤더 정보 중 사용하지 않는 확장 헤더 데이터 (Extension Header Data: 32 bit)부분에 PLC를 위한 추가정보를 전송하는 방법을 제안한다. 이렇게 함으로써 송신단 기반의 알고리즘이 갖는 호환성 문제를 해결하고 음질 개선 성능을 향상시킬 수 있다. 추가적으로 발생하는 지연 (delay)는 이미 수신단에서 네트워크상의 지연을 조정하기 위해 지터 버퍼 (jitter buffer)를 갖고 있기 때문에 제안하는 알고리즘으로 인해 발생하는 추가 지연은 없다. G.729 PLC를 위한 추가 정보는 LP 파라미터 합성용 MA필터 인덱스, 여기신호, 여기신호 이득 및 잔여신호 이득 파라미터로 프레임당 16 bit를 할당한다. 이는 RTP payload 전송 시 음성 데이터를 두 프레임인 20 ms 단위로 전송하기 때문이다. 성능 평가 결과 기존 대비 13.5%의 성능 향상을 보였다.

핵심용어: G. 729 PLC, VoIP, 패킷 손실 개선, RTP

투고분야: 음성처리 분야 (2)

It has a falling off in speech quality, especially when consecutive packet loss occurs, even if a vocoder implemented in the packet network has its own packet loss concealment (PLC) algorithm. PLC algorithm is divided into transmitter and receiver algorithm. Algorithm in the transmitter gives superior quality by additional information, however, it is impossible to provide mutual compatibility and it occurs extra delay and transmission rate. The method applied in the receiver does not require additional delay. However, it sets limits to improve the speech quality. In this paper, we propose a new method that puts extra information for PLC in a part of Extension Header Data which is not used in RTP Header. It can solve the problem and obtain enhanced speech quality. There is no extra delay occurred by the proposed algorithm because there is a jitter buffer to adjust network delay in a receiver. Extra information, 16 bits each frame for G.729 PLC, is allocated for MA filter index in LP synthesis, excitation signal, excitation signal gain and residual gain reconstruction. It is because a transmitter sends speech data each 20 ms when it transfers RTP payload. As a result, the proposed method shows superior performance about 13.5%.

Keywords: G.729 PLC, VoIP, Packet Loss Improvement, RTP

ASK subject classification: Speech Signal Processing (2)

I. 서론

패킷 교환방식을 사용하는 유/무선 서비스가 널리 대중화되면서 기존의 회선교환망을 이용한 통신서비스대신 휴대폰이나 무선인터넷과 같은 다양한 멀티미디어 서비스가 빠르게 확산되고 있다. 그 이유는 기존 회선교환망에 비해 저렴한 이용료로 다양한 서비스를 제공받을 수 있기 때문이다. 패킷 교환망을 이용한 다양한 서비스 중 대표적인 음성통신 서비스는 휴대전화 서비스와 인터넷 전화를 들 수 있다. 이들 서비스는 PSTN을 이용하는 기존의 전화와는 달리 패킷 교환 방식을 사용한다. 이는 전송 시 다양한 형태의 패킷 손실이 발생하는데, 패킷 교환 망에서 발생하는 패킷 손실의 종류는 지연 (delay), 지터 (jitter), 손실 (loss) 등이 있다. 이 중 지연이나 지터의 경우에는 단일 프레임 단위로 발생하지만 손실의 경우에는 연속적으로 발생하는 경우가 대부분이다. 이로 인해 음성 통신의 품질을 보장받지 못하게 된다 [1-3]. 이러한 패킷 손실로 인한 음질 저하를 막기 위해 패킷 네트워크망에 사용되는 음성부호화기들의 대부분은 자체적인 PLC 알고리즘을 보유하고 있다. PLC 알고리즘은 크게 송신단 기반의 방법과 수신단 기반의 방법으로 나뉘어진다. 송신단 기반의 방법은 손실이 발생할 것에 대비해 부가 정보를 보내는 것으로 추가 bit할당을 해야 하며 이로 인해 발생하는 지연도 감수해야 한다. 대표적인 송신단 기반의 방법으로 FEC (Forward Error Correction)이 있다 [4][5]. 이에 반해 수신단 기반의 방법은 별도의 추가 정보 없이 이전에 수신한 데이터에 의존해 손실된 데이터를 복원하는 방법으로 현재 패킷 네트워크망에서 사용하고 있는 방법이다 [6-9]. G.729와 같은 CELP (Code Excited Linear Prediction)방식의 음성부호화기는 예측기법을 사용하기 때문에 손실로 인한 예러가 그 다음 프레임까지 전파되는 문제점이 있다. 이를 보완하기 위해 많은 연구가 있었지만 송신단 기반의 알고리즘을 이용한 기존의 연구는 현재 사용하고 있는 서비스와 호환이 불가능하기에 실제 적용하기가 어렵고, 수신단기반의 알고리즘으로는 실제 네트워크를 통해 전송되는 RTP 패킷조건에 대한 고려 없이 RTP 패킷 단위를 G.729의 프레임 단위로 동일하게 한 조건의 손실에 대해서만 고려하고 있어 서비스 사업자가 요구하는 음성 품질을 만족하지 못하고 있다 [12]. 본 논문에서는 실제 서비스 중인 인터넷 전화 서비스 망의 상태를 분석하고 이에 맞는 PLC 알고리즘을 제안한다. 연속으로 손실된 프레임 복원을 위해 개선된 G.729 PLC 알고리즘 [10]을 사용하고

추가정보를 이용한 복원 알고리즘을 적용하여 실제 패킷 교환망의 서비스 조건에 맞는 음성 품질을 유지할 수 있다. 본 논문의 2장에서는 인터넷 전화 서비스의 음성 품질 조건과 실제 인터넷에서 발생하는 프레임 손실에 대해 설명하고 3장에서는 G.711과 G.729 음성 부호화기의 PLC 알고리즘에 대해서 설명한다. 4장에서는 연속적인 프레임 손실에 대한 G.729 PLC알고리즘을 설명하고 5장에서는 제안한 알고리즘을 설명한다. 6장에서 실험을 통해 제안한 알고리즘의 성능을 분석하고 7장에서 결론을 맺는다.

II. 인터넷 전화의 음성 품질

현재 인터넷 전화 서비스는 기존의 인터넷 서비스 회사들과 지역 케이블 TV 사업자들이 제공하고 있다. 서비스 형태는 여러 종류가 있고 각각의 특성에 따른 조건들이 있지만 공통적으로는 지연, 음질, 지원하는 음성부호화기의 종류 등이다. 표 1은 인터넷 전화 서비스의 음성 품질 조건이다.

여기서 지연은 종간단의 지연 (End-to-End delay)를 말한다. 음질의 경우 PESQ나 PAMS 또는 E-model으로 측정한다. 음질 측정은 업체 마다 조금씩 차이는 있는데 네트워크 시뮬레이터를 이용하여 표 1의 조건에서 측정하거나 아니면 실제 서비스 망에서 측정한다. 패킷 손실을 2%로 정한 이유는 서비스 사업자가 정한 요구 조건으로 실제 서비스 망에서 발생하는 순수한 패킷 손실이 평균 1.5~1.8% 사이에서 발생하기 때문이다 [12]. 여기에 지터 조건을 동시에 주어서 실제 손실은 더 발생할 수 있는데 지터로 인한 손실은 음성부호화기가 아닌 RTP layer에서 지터버퍼 컨트롤을 이용하여 손실을 최소화한다 [13]. 음질 측정에 사용하는 장비는 Agilent사의

표 1. 음성 품질 조건 [11]
Table 1. Condition of voice quality.

종류		조건
지연 (delay)		< 150ms
음질 (MOS)	normal	G.711 : 4.1 이상 G.729 : 3.8 이상 G.723.1 : 3.9 이상
	jitter(40ms) loss 2%	G.711 : 3.6 이상 G.729 : 3.1 이상 G.723.1 : 3.1 이상
frame size		G.711, G.729 : 20ms G.723.1 : 30ms
반복 테스트 횟수		단시간 200회 (40분) 장시간 1440회 (8시간)

VQT 나 Head Acoustic사의 HMS 장비를 사용하며 네트워크 시뮬레이터는 packet storm이나 shuna storm를 이용한다.

III. G.711 & G.729 PLC 알고리즘

3.1. G.711 PLC 알고리즘

G.711 PLC 알고리즘은 과거 데이터를 저장하고 있다 가 손실이 발생할 경우 과거 저장된 데이터에서 피치정보를 구한 뒤 피치 구간의 1/4 구간을 삼각창함수를 이용하여 OLA (Over Lap Add)를 사용하여 피치 구간을 반복하여 손실된 데이터를 복원한다. 연속적으로 손실이 발생할 경우에는 복원된 데이터에 이득을 20% 감소시킨다. 연속적으로 5프레임 이상 손실이 발생할 경우에는 이전 프레임간의 상관관계가 없다고 판단하여 6번째 프레임부터는 zero padding을 하여 데이터를 생성한다 [7].

G.711 PLC 알고리즘에 대한 블록도를 그림 2에 나타냈다. 손실이 발생하지 않았을 경우에는 히스토리 버퍼에 390샘플 (48.5 ms)의 크기로 저장되며, 손실이 발생할 경우 저장된 히스토리 버퍼로부터 피치 정보를 구한 후 OLA를 통해 피치구간단위로 복원하여 손실된 데이터를 복구한다. 이렇게 복구된 데이터는 다시 히스토리버퍼에 저장된다.

그림 3에서는 연속적으로 손실이 발생할 경우에 대한

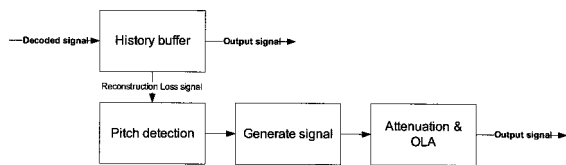


그림 1. G.711 PLC 알고리즘 블록도
Fig. 1. G.711 PLC Algorithm Block Diagram.

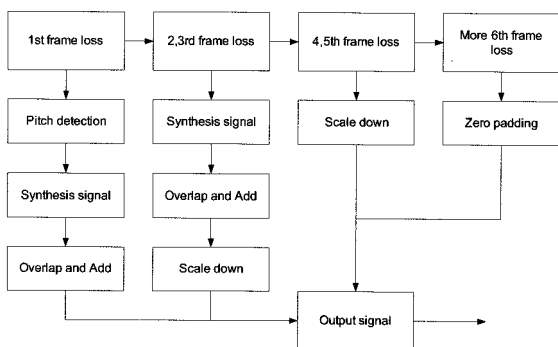


그림 2. 연속 프레임 손실에 대한 G.711 PLC 알고리즘 블록도
Fig. 2. G.711 PLC Algorithm Block Diagram about consecutive frame loss.

블록도를 나타냈다.

한 프레임 이상 연속적으로 손실이 발생할 경우 그림 3과 같이 동작한다. 첫 번째 프레임 손실 복원에서는 히스토리버퍼의 끝에서 1.25 피치 되는 구간에서부터 1피치 구간까지 OLA를 적용하여 손실된 데이터를 복구한다. 2,3번째 프레임 손실 복원에서는 각각 두 배, 세배의 피치 주기를 이용하여 데이터를 복구한 후 이득감쇄를 통해 이득을 낮춘다. 4,5번째 프레임 복구는 앞서 구한 데이터에 이득감쇄만 수행한다. 이후 6번째 이상의 프레임 손실이 발생할 경우에는 모두 0으로 채워 출력한다.

3.2. G.729 PLC 알고리즘

기존의 G.729 PLC 알고리즘은 LP 파라미터 복원 및 여기신호 복원, 이득 복원 이렇게 세 부분으로 나뉘어져 있다.

첫 번째 LP 파라미터 복원은 이전 프레임의 LP 파라미터를 그대로 사용한다. G.729에서는 부호화 시 4차의 MA 예측 필터를 사용하는 데 2개의 MA 예측필터 중 WMSE (Weighted Mean Square Error)를 최소화 하는 MA 필터를 적용하고 그 인덱스를 전송한다. 손실이 발생하여 사용한 MA 필터의 인덱스를 모를 경우 전 프레임에서 사용한 MA 필터를 그대로 사용하여 복원 시 오차가 발생하고 연속적인 프레임 손실이 발생하면 그 오차는 더욱 커지게 된다 [8].

두 번째 여기신호는 복원 이전 프레임의 피치 주기를 이용하여 피치 주기를 복원하고 이를 바탕으로 여기신호를 생성한다. 피치 주기 복원은 식 (1)을 이용한다.

$$L_f = \text{MIN}(L_p + 1, L_{\text{max}}), \quad L_{\text{max}} = 145 \quad (1)$$

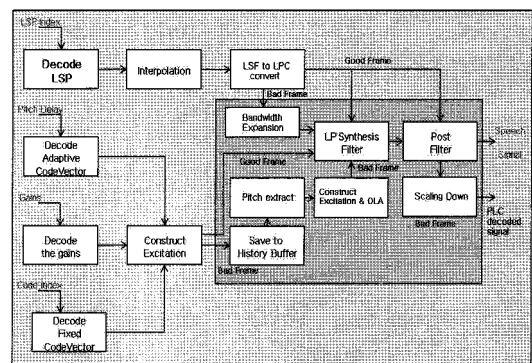


그림 3. 연속적인 패킷 손실 복원 알고리즘의 블록 다이어그램
Fig. 3. Block Diagram of consecutive packet loss concealment Algorithm.

복원된 피치주기 L_f 는 과거 프레임의 피치주기 L_p 에 1씩 증가시킨 후 사용하며 복원된 피치주기가 피치주기 최대값 L_{max} 보다 큰 경우에는 최대값을 사용한다.

고정 코드북의 인덱스와 부호는 각각의 13bit와 4bit 파라미터를 랜덤 함수를 통해 구한 뒤 복원한다.

세 번째 이득 복원 알고리즘은 식 (2)와 같다. 적응 코드북 이득 g_p 과 고정 코드북 이득 g_c 는 이전 프레임의 각각의 이득을 감쇠시켜 얻는다.

$$\begin{aligned} g_p &= \text{MIN}(g_{p_prev} \times 0.9, 0.9) \\ g_c &= g_{c_prev} \times 0.98 \end{aligned} \quad (2)$$

IV. 수신단 기반의 연속적인 패킷 손실에 대한 G.729 PLC 알고리즘

수신단 기반의 연속적인 패킷 손실에 대한 G.729 PLC 알고리즘은 그림 4와 같다. 과거 버퍼의 데이터를 가지고 손실 패킷을 복원하는 방법은 4.1, 4.2, 4.3에 기술하였으며, 그 내용은 각각 LP (Linear Prediction) 파라미터 보간, 여기신호 복원, 여기신호이득 복원 방법이다.

손실 없이 수신된 음성 프레임 복호화 과정에서 생성된 여기원 데이터를 별도의 버퍼 (history buffer)에 저장한다. 이후 음성프레임의 손실이 발생할 경우 앞서 저장된 버퍼로부터 pitch 정보와 여기원 파형을 예측한 후 파형의 연속성을 보존하기 위해 history buffer의 데이터와 보간 (interpolation)을 통해 새로운 여기원을 생성한다. 이렇게 생성된 여기원과 LP 파라미터 보간을 통해 생성된 LP 파라미터와 합성하여 손실된 음성 데이터를 복원한다.

4.1. LP (Linear Prediction) 파라미터 보간

이전 프레임의 LP 파라미터를 그대로 사용하지 않고 보간을 통한 대역폭 확장 방법을 적용하였다. LP (Linear Prediction) 합성 필터의 대역폭을 확장한다는 것은 필터의 피크를 약간씩 줄이고 대역폭을 조금 늘리는 것을 의미한다.

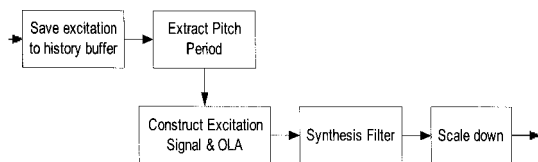


그림 4. 여기신호 복원 방법
Fig. 4. Excitation signal recovery method.

$$A_{PLC}(z) = 1 - \sum_{j=1}^{10} \beta \alpha_j z^{-j} \quad (3)$$

식 (3)를 적용함으로써 이전 프레임으로부터 구해진 LP 합성필터는 조금 더 부드러운 형태가 된다. β 는 0.98로 고정된 값을 사용한다 [12].

이로 인해 천천히 변화하는 음성신호의 스펙트럼 특성을 반영할 수 있고, 합성음성을 보다 부드러운 음성 변화 구간을 얻을 수 있다. 그러나 G.729에서는 2개의 MA 예측필터를 사용하여 이중 WMSE (Weighted Mean Square Error)가 작은 필터를 선택하여 인덱스를 함께 전송한다. 수신측에서는 데이터 손실로 인해 MA 예측 필터의 인덱스를 알 수 없기 때문에 이전 프레임에서 사용한 MA 예측필터를 그대로 사용한다. 송신단에서 사용한 필터와 수신단에서 사용한 필터가 틀리게 되면 오차는 커지게 된다.

4.2. 여기신호 복원

G.711 PLC의 알고리즘을 적용하여 손실된 패킷의 여기신호를 복원한다. 우선 별도의 버퍼 (history buffer)를 두어 정상적으로 수신된 이전 프레임의 여기신호를 저장한다. 이후 손실된 패킷의 피치주기를 복원할 때 이전 프레임의 피치 주기를 이용하여 히스토리버퍼에서 검색 구간을 20샘플로 결정하고 교차상관도를 구한뒤 이를 이용하여 손실된 패킷의 피치 주기를 구한다. 식 (4)에 교차상관도를 이용한 피치 검색방법을 나타냈다 [12].

$$C(t) = \frac{\left(\sum_{n=0}^{19} e[n] \cdot e[n-t] \right)^2}{\sum_{n=0}^{19} e[n-t] \cdot e[n-t]} \quad , \quad p-2 < t < p+2 \quad (4)$$

여기서 p는 이전 프레임의 피치주기이고 $e[n]$ 은 히스토리버퍼에 저장된 여기신호이다. 이전 피치주기의 전후 2씩의 범위에서 상관도 값이 가장 큰 자연 1을 피치주기로 정한다. 구해진 피치 주기를 이용하여 OLA (OverLap Addition)한다. 손실 패킷의 여기신호의 생성은 찾은 피치를 기반으로 피치주기를 반복하여 사용한다. 그림 3에서 여기 신호 합성 방법에 대한 구성은 아래의 그림 4와 같다.

4.3. 여기신호 이득 복원

기존 G.729 embedded PLC에서는 적응, 고정 코드북

이득을 감쇠시킨다. 재귀 구조를 갖는 G,729 의 특성상 여기원의 이득 감쇠가 다음 프레임에 영향을 끼치게 되어 PLC로 인한 오차가 이후까지 전파되는 결과를 낳는다. 이런 영향을 없애기 위해 첫 번째 PLC 상황에서의 합성에서는 감쇄를 하지 않고 이전 프레임의 이득을 그대로 적용한다. 이후 연속적으로 패킷 손실이 발생할 경우 합성된 전체 신호에 대해서 이득을 10%씩 감쇄하여 사용한다.

$$s(n) = G_{PLC} \cdot s(n) \quad (5)$$

$$G_{PLC, Current} = 0.9 \cdot G_{PLC, Previous} \quad (6)$$

식 (5)는 프레임손실이 연속적으로 발생했을 경우 두 번째 프레임부터 적용하는 전체 이득 조절이다. 식 (6)과 같이 연속적으로 프레임 손실이 발생할 경우 이득을 갱신한다.

V. 제안한 알고리즘

상용화된 VoIP 서비스 환경을 분석하면 G,729의 경우 RTP 패킷 당 2 프레임의 음성 데이터를 전송하는 것이 일반적이다 [11]. 때문에 음성 데이터의 손실은 최소 2 프레임이상 연속적으로 발생하게 된다.

본 논문에서는 패킷 손실이 연속적으로 발생한다는 조건에서 음질열화를 최소화하는 알고리즘을 제안한다. 수신단 기반의 알고리즘으로는 음질개선에 한계가 존재하기 때문에 기존 시스템과의 호환이 가능한 송신단 기반의 알고리즘을 제안한다.

RTP payload는 그림 5와 같이 구성되어있다. 그 중 header extension 부분은 추후 사용 가능성을 위해 예비된 부분으로 현재는 사용하고 있지 않으며 사용 시에는 x=1로 하여 사용한다. 수신측에서는 x bit를 보고 header extension의 데이터를 읽어서 처리하는데 서로 약속된 데이터가 아니면 무시하도록 되어 있다. 본 논문에서는

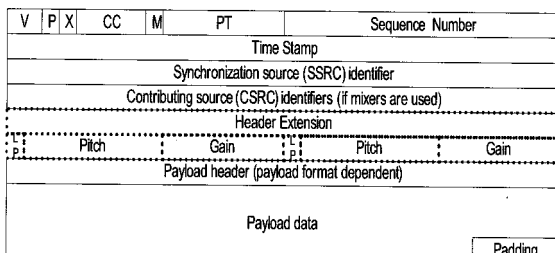


그림 5. RTP payload 형식
Fig. 5. RTP payload format.

header extension에 32 bit를 할당하여 한 프레임당 16 bit씩 두 프레임에 대한 추가정보를 산도록 하였다. 이는 RTP payload 하나당 G,729 두 프레임 (20 ms) 데이터를 전송하기 때문이다.

G,729와 같은 CELP(Code Excited Linear Prediction) 형태의 음성부호화기에서 음질에 가장 영향을 미치는 파라미터는 LP 파라미터와 피치 주기, 이득값이다. 표 2에 패킷 손실 위치에 따른 파라미터의 중요도를 나타냈다. 유성음 구간이나 음성이 시작되는 전이구간에서는 피치 정보가 가장 중요하고 묵음이나 음성이 끝나는 전이구간에서는 이득이 가장 중요하다 [12].

세 가지 파라미터를 16 bit로 양자화 하여 RTP header extension 위치에 추가하여 기존 RTP 통신과의 호환성을 유지하도록 하였다. 각 파라미터에 할당한 bit는 LP 파라미터 1 bit, 피치주기 정보에 8 bit, 이득 정보에 7 bit로 프레임당 총 16 bit를 할당한다. 5.1, 5.2, 5.3에서 각각 LP 파라미터, 피치 주기, 이득값에 대한 양자화 및 복원에 대해 기술한다.

그림 6은 부가정보를 이용한 손실된 프레임 복원 방식을 나타내었다.

패킷 손실이 발생하면 음성 데이터는 두 프레임 데이터의 손실이 발생한다. 이를 복원하는 방법은 지터버퍼에 있는 다음에 복호화할 패킷에서 RTP header에 있는 부가 정보와 이전 프레임의 파라미터 정보를 이용하여 손실된

표 2. 패킷 손실 위치에 따른 파라미터 중요도
Table 2. Priority in location of Packet Loss.

패킷손실 상황	중요 순서
묵음	Gain > LSF > fixed-codebook > Pitch
음성 → 묵음	Gain > LSF > fixed-codebook > Pitch
묵음 → 음성	Pitch > fixed-codebook > LSF > Gain
음성	Pitch > fixed-codebook > LSF > Gain

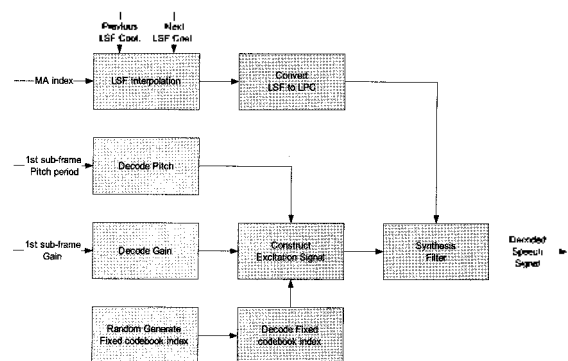


그림 6. 부가정보를 이용한 손실된 패킷 복원
Fig. 6. Lost packet reconstruction by using Additional Information.

데이터를 복원한다. 만일 연속적으로 두 개 이상의 RTP 패킷이 손실될 경우에는 마지막 손실된 패킷은 그 다음 정상적으로 수신된 패킷의 부가정보를 이용하여 복원이 가능하지만 이전 패킷은 부가정보까지 손실되어 그림 6 과 같은 방법으로는 복원이 불가능하다. 이 경우에는 이전 프레임의 데이터만을 이용하여 데이터를 복원한다. 그림 7에는 연속적으로 두 개 이상의 RTP 패킷이 손실되었을 경우에 대한 복원 방법을 나타내었다.

5.1. LP 파라미터 양자화 및 복원

LP 파라미터는 한 프레임당 10자의 계수를 벡터 양자화하며 전체 데이터 중 22.5%를 차지할 정도로 그 중요도가 매우 높다. G.729에서는 부호화 시 2개의 4차 MA 예측 필터를 이용하여 WMSB를 최소화 하는 필터를 사용한다. 분석 결과 2개의 MA 예측 필터 중 사용한 필터를 알 수 있다면 이전 프레임에서 사용한 LP 파라미터와 지터버퍼에 저장된 미래 프레임의 LP 파라미터를 보간하여 적용해도 충분히 손실된 데이터를 복원이 가능하다. 따라서 LP 파라미터 18 bit를 전부 보내지 않고 MA 예측 필터 인덱스만 전송하여도 충분하다.

보통 수신측에서는 지터현상으로 인한 패킷 손실을 방지하게 위해 일정크기의 버퍼를 보유하고 있다. 이 지터 버퍼에 저장된 데이터를 이용하여 식 (7)과 같이 보간을 통해서 손실된 패킷에서의 LP 파라미터를 복원한다.

$$\hat{\Omega}_{n_PLC}(w) = 0.5 \cdot \hat{\Omega}_{n-1}(w) + 0.5 \cdot \hat{\Omega}_{n+1}(w) \quad (7)$$

식 (1)에서 $\hat{\Omega}_{n_PLC}(w)$ 는 손실된 n번째 LSF 계수이고 $\hat{\Omega}_{n-1}(w)$ 는 복호화된 이전 프레임의 LSF 계수이다. $\hat{\Omega}_{n+1}(w)$ 는 지터버퍼에 다음 프레임의 LSF 계수이다.

표 3는 각 파라미터에 대해 각각 하나씩만 복원하였을 경우 음질 개선 효과에 대해 나타냈다. LP 파라미터의 경우 인덱스 정보만 전송하여도 기존의 방법에 음질 개선

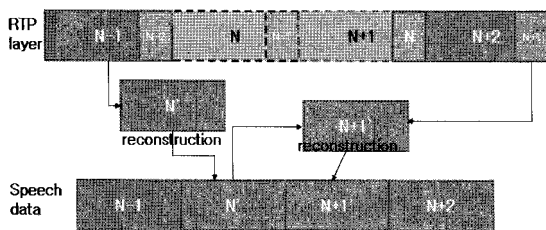


그림 7. 연속적인 패킷 손실 상황에서의 손실된 패킷 복원
Fig. 7. Lost packet reconstruction in Consecutive Packet Loss Situation.

효과가 있음을 알 수 있다.

5.2. 피치 파라미터 양자화 및 복원

피치정보는 부 프레임 단위로 양자화되고 전송된다. 첫 번째 부 프레임에서는 8 bit를 할당하고 두 번째 부 프레임에서는 첫 번째 부 프레임의 피치 주기와의 차분을 이용하여 5 bit만 전송한다. 본 논문에서는 첫 번째 부 프레임의 피치 정보 8 bit를 그대로 전송하고 두 번째 부 프레임의 피치는 식 (7)과 같이 첫 번째 부프레임의 피치 주기의 주변 범위에서 상관관계를 이용하여 구한다. 그림 7과 같이 연속적인 패킷 손실로 인해 부가정보를 사용할 수 없을 경우에는 이전 프레임의 데이터를 이용하는 4.2절의 방법을 사용하여 복원한다.

$$C(l) = \frac{\left(\sum_{n=0}^{39} s[n] \cdot s[n-l] \right)^2}{\sum_{n=0}^{39} s[n-l] \cdot s[n-l]}, \quad p-2 < l < p+2 \quad (7)$$

5.3. 이득 파라미터 양자화 및 복원

표 2에서처럼 목음이나 유성음에서 북음으로 바뀌는 전이구간에서는 이득 값이 가장 중요한 영향을 미친다. 첫 부프레임의 이득 값 7 bit를 부가정보로 전송하고 두 번째 부 프레임은 앞 또는 뒤의 정상 프레임 데이터와 보간하여 사용한다. 부가 정보가 없는 경우에는 4.2절과 같이 수신된 데이터를 기반으로 복원한다.

VI. 실험 및 결과

제한한 알고리즘의 성능평가를 위해 C-언어를 사용하여 PC에 구현하였다. 성능평가를 위해 사용한 시료는 대표적인 분장 남,녀 각각 5명이 5번씩 발생하여 시료로 사용하였다. 음성 시료는 다음과 같다.

표 3. 발생 1에 대해 각 파라미터별 패킷 손실 은닉 알고리즘의 MOS값 [10]

Table 3. Packet Loss Concealment Algorithm Result.

패킷 손실 조건		jitter 40ms / loss 2%	
기존의 G.729 PLC		2.81 / 3.1	
제한한 알고리즘	LP 복원	2.99/3.2	
	여기신호 복원	3.04/3.42	
	여기신호이득 복원	2.89/3.15	
	종합	3.18 / 3.68	

- 발성 1: /인수네 꼬마는 천재소년을 좋아한다./
- 발성 2: /예수님께서 천지창조의 교훈을 말씀하셨다./
- 발성 3: /창공을 헤쳐나가는 인간의 도전은 끝이 없다./
- 발성 4: /숭실대학교 정보통신공학과 음성통신연구팀이다./
- 발성 5: /공일이삼사오육칠팔구/

음성자료 조건에 잡음을 첨가하지 않은 이유는 모든 VoIP 단말은 전처리/후처리 과정을 통해 유입되는 잡음을 제거하기 때문에 잡음 세기에 따른 음질측정은 측정 조건에서 제외하였다.

실험 조건은 실제 상용서비스망에서 사용하는 조건과 동일하게 2프레임/패킷으로 적용하였으며, 패킷당 프레임수가 음질에 영향을 미치는 바를 확인하기 위해 비교 대상으로 1프레임/패킷 조건을 추가하였다. FER (Frame Error Rate)도 2%로 고정하였다.

각 발성 당 40회씩 반복하여 패킷 손실을 발생시킨 후 그 중 최소값과 평균값을 나타내었다.

표 4는 본 논문에서 제안한 알고리즘과 기존의 G.729 PLC 알고리즘의 성능을 각각 보여준다. RTP패킷 당 1프레임을 기준으로 할 경우에는 기존의 알고리즘도 서비스 사업자가 요구하는 기준 MOS 3.1에 근접하지만 실제 서비스 기준인 RTP패킷당 2프레임으로 적용할 경우 음질이 현저히 낮아지는 것을 알 수 있다. 제안한 알고리즘의 성능향상의 원인은 첫 번째로 부가정보를 전송함으로써 CELP 형태의 음성부호화기에서 손실 오차가 이후 프레임에 영향을 미치는 것을 방지하였기 때문이다. 두 번째는 부가정보를 이용하지 못할 경우의 손실 은닉 알고리즘을 기본적으로 2 프레임이상 연속 손실에 대해 최적화 되도록 설정했기 때문이다.

표 4. 제안한 패킷 손실 은닉 알고리즘의 MOS값
Table 4. Proposed Packet Loss Concealment Algorithm Result.

	G.729 PLC				Proposed PLC			
	1 frame /packet		2 frame /packet		1 frame /packet		2 frame /packet	
	MIN	AVG	MIN	AVG	MIN	AVG	MIN	AVG
발성1	3.08	3.63	2.81	3.1	3.26	3.72	3.18	3.68
발성2	3.23	3.67	2.90	3.14	3.38	3.77	3.27	3.71
발성3	3.18	3.6	2.72	3.08	3.31	3.73	3.16	3.6
발성4	3.01	3.62	2.83	3.22	3.28	3.7	3.19	3.61
발성5	3.35	3.68	2.79	3.03	3.26	3.69	3.15	3.63
평균	3.17	3.64	2.81	3.114	3.298	3.72	3.19	3.646

VII. 결론

패킷교환 망에서는 단말간의 지연 조정, 패킷 지연, 지터, 패킷 손실 등 다양한 이유로 좋은 음질을 보장 받지 못하기 때문에 이런 손실을 보상하는 방법에 대한 연구가 필수적이다. 기존의 제안된 연구들을 보면 실험 조건이 네트워크로 전송하는 패킷 당 하나의 음성 프레임을 전송하는 것으로 가정한다. 그러나 실제 서비스망에서는 네트워크의 효율적인 사용을 위해 패킷당 두 개의 음성프레임을 전송한다. 이렇게 함으로써 패킷의 오버헤드를 줄일 수 있기 때문이다. 이로 인해 네트워크 효율은 개선되었지만 전체적인 지연이 증가하고, 손실이 발생할 경우 음성 데이터는 두 프레임씩 손실이 발생하여 음질이 급격히 저하되는 단점이 있다. 이에 본 논문에서는 연속적인 프레임 손실 상황에 적합한 송신단기반의 CELP형 음성 부호화기의 PLC 알고리즘을 새로이 제안하였다. RTP header 데이터 중 사용하고 있지 않는 extension header 부분에 손실 복원을 위한 부가정보로 LP 파라미터, 피치 주기, 피치이득 및 잔여신호 이득을 양자화하여 전송함으로써 손실로 인한 오차가 이후 프레임까지 전파되는 것을 최소화하였고 패킷당 두 프레임의 부가정보를 전송하도록 하여 손실을 복원할 수 있도록 하였다.

제안한 알고리즘은 기존의 방법에 비해 성능을 향상시켰지만 burst packet loss의 경우 마지막 두 프레임만 부가정보를 이용하고 나머지 손실 프레임은 수신된 데이터를 기반으로 은닉 알고리즘을 사용할 수밖에 없었다. 이에 대해 부가정보를 더욱 작은 bit로 양자화하여 더 많은 프레임의 부가정보를 전송하는 연구가 필요하다.

참고 문헌

1. J. Bolot and A. Vega Garcia, "Control Mechanisms for Packet Audio in the Internet," in *Proceeding of IEEE INFOCOM*, vol.1, no.1 pp. 232-239, March 1996.
2. C. Perkins, O. Hodson and V. Hardman, "A Survey of Packet Loss Recovery Techniques for streaming Audio," *IEEE Network*, Vol. 12, no. 5, pp.40-48, Sep. 1998.
3. C. Romer, M. Podolsky and S. McCanne, "Simulation of FEC-based error control for Packet Audio on the Internet," in *Proceedings of IEEE INFOCOM*, vol.2, no.1, pp. 505-515, April, 1998.
4. T. J. Kostas, et al, "Real-time voice over packet-switched networks," *IEEE Network*, vol. 12, no.1, pp. 18-27, Jan.-Feb. 1998
5. J. C. Bolot, et al, "Adaptive FEC-based Error Control for Internet Telephony," *Proceeding INFOCOM 99*, vol. 3, pp. 21-25, March 1999.

6. ITU-T Rec. G.711 Appendix I, *A high quality low-complexity algorithm for packet loss concealment with G.711*, Sep. 1999.
7. ITU-T Rec. G.729, *Coding of speech at 8kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)*, March, 1996
8. ITU-T Rec. G.723.1, *Dual rate speech coder for multimedia communication transmitting at 5.3 and 6.3 kbit/s*, Mar. 1996
9. ETSI TS 126.091, *AMR speech Codec; Error concealment of lost frames*, 2009-01
10. 홍성훈, 김중국, 배명진, "VoIP 시스템의 음질개선," 2009년도 음향학회 음성통신 및 신호처리 학술대회 논문집, 26권 1호 pp.49-52
11. VoIP 음성 품질 평가 항목, 하나로텔레콤, 2007
12. 윤성원, 강홍구, 윤대희, "음성신호의 특성을 고려한 패킷 손실 은닉 알고리즘," *한국통신학회 논문지*, vol.31, No.7c, 2006년 7월
13. N. Kitawaki and K. Itoh, "Pure Delay Effects on Speech Quality in Telecommunications," *IEEE Journal on Selected Areas in Communication*, Vol.9, No.4, pp.586-593, May 1991

저자 약력

•홍성훈 (Seong-Hoon Hong)



1997년 2월 : 송실대학교 정보통신공학과 (공학사)
 1999년 2월 : 송실대학교 정보통신공학과 (공학석사)
 1999년 2월~현재 : C&S Technology IP Telephony
 사업부 수석연구원
 2004년 3월~현재 : 송실대학교 정보통신공학과 박사
 과정

•배명진 (Myung-Jin Bae)



현재 송실대학교 정보통신전자공학부 교수
 한국음향학회지 제 26권 제4호 참조