

음성 특성을 이용한 G.711 패킷 손실 은닉 알고리즘의 성능개선

한승호(ICU), 김진술(ETRI), 이현우(ETRI), 류원(ETRI),
한민수(ICU)

<차 례>

- | | |
|-----------------------------------|----------------|
| 1. 서론 | 4. 성능 평가 실험 |
| 2. G.711의 패킷 손실 은닉 알고리즘 | 4.1. 패킷 손실 모델링 |
| 3. 음성특성을 이용한 패킷 손실 은닉 알고리즘 | 4.2. 성능 평가 방법 |
| 3.1. 유/무성음 분류에 따른 적응적 손실 양단 정보 이용 | 5. 평가 및 결과 분석 |
| 3.2. 적응형 이득 조절 | 6. 결론 |

<Abstract>

Performance Improvement of Packet Loss Concealment Algorithm in G.711 Using Speech Characteristics

Seung Ho Han, Jinsul Kim, Hyun-Woo Lee, Won Ryu, Minsoo Hahn

Because a packet loss brings about degradation of speech quality, VoIP speech coders have PLC (Packet Loss Concealment) mechanism. G.711, which is a mandatory VoIP speech coder, also has the PLC algorithm based on pitch period replication. However, it is not robust to burst losses. Thus, we propose two methods to improve the performance of the original PLC algorithm in G.711. One adaptively utilizes voiced/unvoiced information of adjacent good frames regarding to the current lost frame. The other is based on adaptive gain control according to energy variation across the frames. We evaluate the performance of the proposed PLC algorithm by measuring a PESQ value under different random and burst packet loss simulating conditions. It is shown from the experiments that the performance of the proposed PLC algorithm outperforms that of PLC employed in ITU-T Recommendation G.711.

1. 서 론

최근 통신 사업자 및 일반 이용자들의 관심이 고조되고 있는 VoIP 서비스는 인터넷과 같은 IP 네트워크를 이용하여 패킷 기반으로 음성 통신을 수행한다. IP 망을 기반으로 음성을 전송함에 따라 IP 네트워크의 근본적인 특성인 지연(delay), 지터(jitter), 네트워크 단전 등에 기인하여 전송된 음성 데이터의 손실이 발생하며, 이러한 수신단에서의 음성 데이터의 손실을 패킷 손실이라 한다.

패킷 손실이 발생하면 수신단에서 음성을 재생할 때 음성 끊김 현상, 부자연스러운 음성 연결 등이 발생하여 통화 품질의 저하를 가져온다. 특히, 연속적인 손실이 발생하거나, 높은 손실률을 가지고 손실이 발생하게 되면 통화 품질에 큰 저하를 유발한다. VoIP에서 음성 데이터 전송은 실시간으로 이뤄지기 때문에 패킷 손실이 발생했을 때 일반 데이터의 전송에서와 같이 손실된 패킷을 재전송하는 것이 불가능하다. 따라서 통화품질을 유지하기 위해서는 패킷 손실이 발생했을 때 이를 복호화 과정에서 보상해 주는 기술이 필요하며, 이를 패킷 손실 은닉(PLC: Packet Loss Concealment) 또는 프레임 손실 은닉(Frame Erasure Concealment) 알고리즘이라 한다. 패킷 손실 은닉 알고리즘을 통해 손실된 음성 데이터를 예측하여 재생함으로써 통화 품질의 큰 저하를 방지할 수 있다.

VoIP에서 주로 사용되는 G.711, G.723.1, G.729A 등의 음성 코덱들은 패킷 손실이 발생했을 때 통화 품질을 유지하기 위해 이를 보상해주는 고유의 패킷 손실 은닉 알고리즘을 내장하고 있다. 각 코덱에서 패킷 손실 은닉 알고리즘은 기본적으로 이전에 수신된 음성 패킷의 정보를 이용하여 손실된 음성 패킷의 정보를 예측하는 방법에 의해 수행된다[1][2][3]. G.711의 경우는 PCM(Pulse Code Modulation) 기반의 코덱이기 때문에 음성데이터를 직접 이용하여 보상을 수행하며, G.723.1, G.729A와 같은 Hybrid 코덱들은 LSF(Line Spectral Frequency)와 같은 음성 파라미터의 예측, 여기 신호의 예측을 통해 보상을 수행하게 된다.

본 연구에서는 G.711에 대해 패킷 손실 은닉 기술의 성능 향상을 위한 방법을 제안한다. 이를 위해 G.711 Appendix I 패킷 손실 은닉 기술의 성능 평가 결과를 바탕으로 음성의 특성을 이용하여 적응적인 손실 양단 정보의 이용과 적응적인 이득 조절 방법에 의한 패킷 손실 은닉 기술의 성능 향상을 위한 방법을 제안한다. 성능평가를 위해 ITU-T에서 제공하는 객관적인 음질 평가방법인 PESQ(Perceptual Evaluation of Speech Quality)를 사용하였다[4]. 또한, 네트워크에서 발생하는 패킷 손실을 모델링하기 위해 ITU-T에서 제공하는 EID(Error Insertion Device)를 이용하여 랜덤(random) 손실과 버스트(burst) 손실을 생성하였다[5].

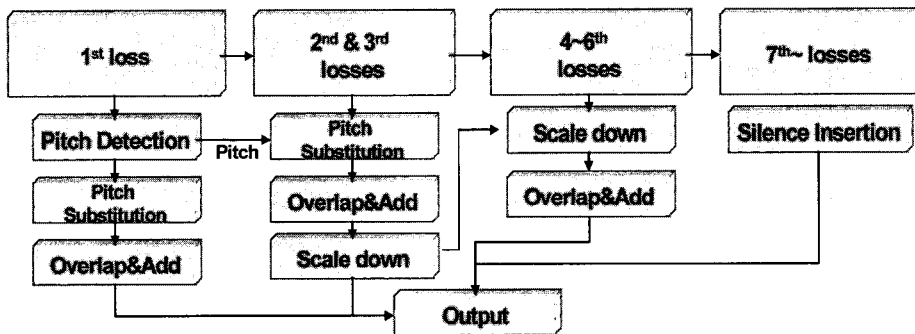
본문의 구성은 다음과 같다. 먼저 2장에서 G.711 Appendix I에 내장되어 있는 패킷 손실 은닉 알고리즘에 대해 소개하고, 3장에서 제안된 음성 특성을 이용한 패킷 손실 은닉 알고리즘을 설명한다. 4장에서 실험 환경 및 평가 방법을 설명하

고, 5장에서는 결과 및 분석을 수행하며, 마지막으로 6장에서 결론을 맺는다.

2. G.711의 패킷 손실 은닉 알고리즘

G.711은 VoIP에서 가장 기본적으로 사용되는 음성 코덱으로 표본화율(sampling rate)은 8 kHz이고, A/μ law PCM 기반의 부호화를 수행하며 64 kbps의 비트율(bit rate)을 가진다. G.711은 샘플단위의 코덱이지만 VoIP에서 이용할 때는 프레임 단위로 음성 부호화를 수행하며 이때 프레임 길이는 10 ms를 사용하거나 구현방법에 따라 변형하여 다른 프레임 길이를 사용할 수 있다[1].

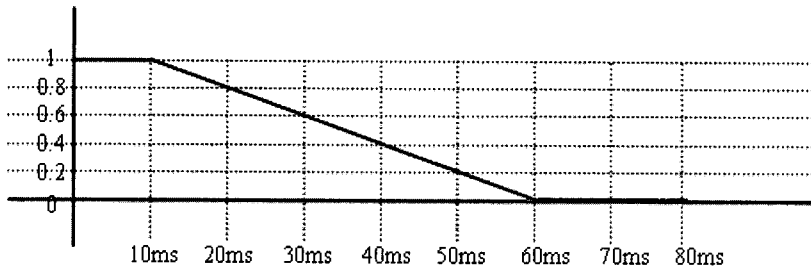
G.711 Appendix I에서 제안하는 패킷 손실 보상 알고리즘은 히스토리(history) 버퍼에 저장되어 있는 48.75 ms의 이전 음성 신호로부터 피치(pitch) 구간을 계산한 후 계산된 피치 구간을 반복해서 삽입하는 피치구간삽입 방법이다[1]. 패킷 손실 보상을 위해 복호화된 출력 신호는 히스토리 버퍼에 저장되며 히스토리 버퍼는 차후 손실된 구간에서의 피치를 계산하고 손실구간 동안의 신호를 추출하는데 사용된다. 만약 패킷 손실이 발생하면 히스토리 버퍼의 내용을 피치(pitch) 버퍼로 복사한 후 피치 버퍼의 음성으로부터 피치를 계산한다. 부드러운 음성 연결을 위해 계산된 피치의 1/4 지점을 삼각창을 이용하여 OLA(OverLap and Add)를 사용하여 피치 구간을 반복하여 삽입한다.



<그림 1> 연속적인 손실에 대한 G.711 PLC 알고리즘 블록도

<그림 1>에서 보는 것과 같이, 연속된 손실에 대해 첫번째 손실에서는 히스토리 버퍼의 내용을 피치 버퍼에 복사한 후 피치 버퍼로부터 피치를 계산한다. 피치 버퍼의 끝에서 1.25 피치 되는 구간에서부터 1 피치 구간까지 OLA 적용하여 피치 버퍼의 끝에서 1.25 피치 구간부터 피치구간을 반복적으로 삽입한다. 2,3 번째 연속 손실에서는 각각 피치 버퍼의 끝에서 2.25 피치, 3.25 피치에서부터 피치 구간을 계산하여 복원을 수행한다. 4~6번째 연속 손실에서는 새로운 피치를 계산하지

않고 이전에 구한 피치 구간을 이용하여 복원을 수행한다. 피치 구간의 반복을 계속할 경우 인공음으로 인한 음질 저하가 발생하므로 손실이 연속으로 발생했을 때는 20%의 감쇠 인자를 이용하여 <그림 2>와 같이 연속된 손실에 대해 이득 조절을 수행하여 복원된 음성의 크기를 계속 감쇠시키며, 7번째 연속 손실부터는 묵음으로 복원한다.



<그림 2> G.711 PLC의 이득 조절

3. 음성특성을 이용한 패킷 손실 은닉 알고리즘

3.1. 유/무성을 분류에 따른 적응적 손실 양단 정보 이용

음성은 시간의 변화에 대해 천천히 변하는 특성을 가지고 있으므로, 손실된 음성 구간과 가장 가까운 구간의 음성이 손실된 음성 구간과 가장 유사한 특징을 보인다. 그러므로 손실 이후의 정보를 이용하여 손실 보상을 수행하게 되면 보다 정확한 손실 보상을 수행 할 수 있다. 그러나 손실 이후의 정보를 사용하게 되면 손실 이후 정보 이용에 따른 부가적인 지연을 유발하게 된다. VoIP에서 지터를 없애기 위해 사용되는 지터 버퍼에는 미래의 음성 정보가 저장되어 있으므로, 부가적인 지연 유발 문제는 지터 버퍼와 연계를 통해 해결할 수 있다.

손실 이후 정보를 이용하기 위해 G.711 Appendix I 패킷 손실 은닉 알고리즘에서 과거의 음성 정보를 저장하는 히스토리 버퍼와 유사하게 손실 이후 정보를 저장하고 있는 미래정보 버퍼를 사용한다. 이 버퍼의 크기는 20 ms로 정의하여 손실 이후 20ms의 음성 정보를 이용한다. 제안한 알고리즘은 만약 패킷 손실이 발생했을 때 미래정보 버퍼에 저장된 손실 이후 정보를 이용할 수 있으면 이를 이용하여 손실을 보상하며, 이용할 수 없을 경우는 기존의 방법과 동일하게 손실을 보상한다. 손실 이후의 정보를 이용할 경우도 손실 이전 정보를 이용할 때와 마찬가지로 피치 구간 반복 알고리즘을 이용하여 음성을 보상한다.

손실 양단의 정보를 이용할 때 음성의 변화 특성을 보장하기 위해 유/무성음

분류에 따른 적응적 손실 양단 정보를 이용한다. 유/무성음 분류기를 통해 손실 이전/이후 음성 프레임의 특성이 상이 할 경우는 손실 양단의 정보를 동시에 사용하여 손실 구간을 복원한다. 이를 위해 히스토리 버퍼와 미래정보 버퍼에 저장되어 있는 손실 이전/이후 정보에 대해 각각 기존에 G.711 Appendix I의 패킷 손실 은닉 방법인 피치 구간 반복 알고리즘으로 복원 신호를 구한다. 각기 구해진 두 가지 종류의 복원 신호를 평균하여 복원 신호로써 이용함으로써 손실 양단 정보를 동시에 사용한 패킷 손실 은닉을 수행하며, 신호의 부드러운 연결을 위해 프레임 연결 부위에서는 기존의 G.711의 패킷 손실 은닉 알고리즘에서와 같이 OLA를 수행한다. 반면에 손실 이전/이후 음성 프레임의 특성이 동일한 경우는 손실 이후 정보만을 이용하여 피치 구간 반복 알고리즘을 적용하여 손실 구간을 복원한다. 따라서 적응적으로 손실 양단 정보를 이용하여 음성의 유/무성음 변이 특성을 보강함으로써 보다 정확한 음성 복원을 수행할 수 있다.

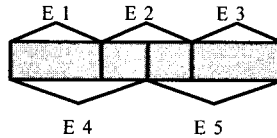
실험에 사용한 유/무성음 분류기는 결정트리 방식으로 분류를 수행한다. 대화체 음성데이터베이스 분석을 통해 CART(Classification And Regression Tree) 알고리즘을 사용하여 결정트리를 생성하였다. CART 알고리즘은 분류 문제에 많이 사용되는 알고리즘으로써 결정 트리를 생성하기 위해 많은 연산이 필요하지만 일단 결정 트리가 생성되면 그 결과는 쉽고 빠르게 이용할 수 있는 장점이 있다[6]. 따라서 실시간 음성 통신에서 큰 지연을 유발하지 않는다. 결정트리를 생성하기 위해 여성화자의 한국어 대화체 음성 1000문장이 사용되었으며, 훈련을 위해 66%의 데이터가 사용되었으며, 나머지 34%의 데이터는 분류 정확도 측정을 위한 실험 데이터로 사용되었다. 설계된 유/무성음 분류기는 훈련 데이터에 대해서는 93.4%, 테스트 데이터에 대해서는 93%의 분류 정확도를 각각 가진다. 이와 같은 방법으로 유/무성음 분류를 수행했을 경우 훈련데이터에 종속적인 결정트리가 만들어지는 단점이 있으므로 최대한 실제 환경과 유사한 훈련데이터를 사용하는 것이 중요하겠다.

결정트리를 생성하기 위해 ZCR(Zero Crossing Rate), 자기 상관 함수 계수, 인트라(intra) 프레임 절대치 합 비율, 인터(inter) 프레임 절대치 합 비율이 특징 벡터로써 사용되었다. ZCR은 신호에서 영점을 교차하는 수를 나타낸다. 자기 상관 함수 계수(η)는 자기상관도함수(ϕ)에서 첫 번째, 두 번째 값 사이의 비율로 식 (1)과 같이 정의된다.

$$\eta = \frac{\phi(1)}{\phi(0)}, \quad \phi(k) = \frac{1}{m} \sum_{n=0}^{m-1} x(n)x(n+k) \quad (1)$$

인트라 프레임 절대치 합 비율은 한 프레임 안에서 신호의 절대 크기 변화율을 나타내며, 인터 프레임 절대치 합 비율은 인접 프레임 사이에서 신호의 절대

크기 변화율을 나타낸다. 인터 프레임 절대치 합과 인트라 프레임 절대치 합을 계산할 때는 식 (2)가 사용되었다. 인트라 프레임 절대치 합 비율은 총 3종류가 계산되며 식 (3)에서 ER1, ER2, ER3으로 표현된다. 한 프레임을 각각 <그림 3>과 같이 E4, E5로 2등분, E1, E2, E3으로 3등분하여 식 (3)과 같이 계산한다. 식에서 E1, E2, E3, E4, E5는 한 음성 프레임안의 구간에서의 절대치 합을 의미한다.



<그림 3> 인트라 프레임 변화율 측정 구간

$$E = \frac{1}{N} \sum_{n=0}^{N-1} |x(n)| \quad (2)$$

$$ER1 = \frac{E1}{E2}, ER2 = \frac{E2}{E3}, ER3 = \frac{E4}{E5} \quad (3)$$

3.2. 적응형 이득 조절

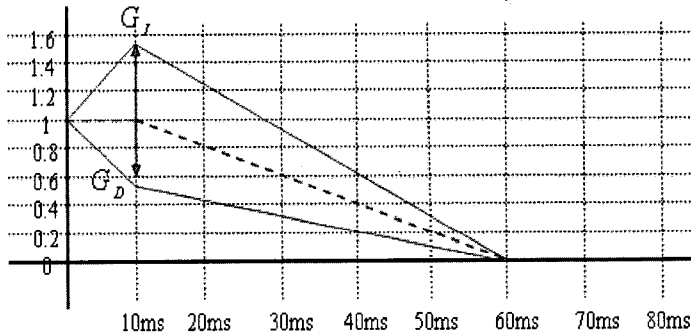
G.711 Appendix I 패킷 손실 은닉 알고리즘은 2장에서 살펴 본 것과 같이 연속적인 프레임 손실에 대해 언제나 동일한 비율로 감쇠하는 이득 조절을 수행한다. 본 연구에서는 음성의 크기 변화 특성을 보다 잘 반영하기 위해 패킷 손실이 발생했을 때 손실 이전 음성 프레임 신호의 절대 크기 변화 특성을 분석하여, 적응적으로 감쇠 이득을 조절하는 알고리즘을 제안한다.

$$G_I = \frac{2 \times E_{i-1} - E_{i-2}}{E_{i-1}} \quad ; E_{i-2} < E_{i-1} \quad (4)$$

$$G_D = 2 - \left(\frac{2 \times E_{i-2} - E_{i-1}}{E_{i-1}} \right) \quad ; E_{i-2} \geq E_{i-1}$$

패킷 손실이 발생했을 때 이전 두 프레임의 절대치 합을 식 (2)으로 각각 계산하여 절대 크기가 증가 혹은 감소하고 있는 경향을 판별하여 식 (4)를 이용해 절대 크기가 증가하고 있던 음성에 대해 초기 이득 값 GI와 절대 크기가 감소하고 있던 음성에 대해 초기 이득 값 GD를 계산한다. 식에서 Ei-1, Ei-2는 각각 손실 이전 두 프레임의 절대치 합이다. GI를 구하는 식은 손실 이전 두개 프레임 신호의 절대치 합을 이용하여 손실된 프레임 신호의 절대치 합을 예측하여 이 값을 이전 프레임 신호의 절대치 합으로 나눠 준 것이다. GI는 항상 1~2사이의 값을 가지게

된다. GD를 구하는 식도 유사하게 E_i-2 를 기준으로 증가하는 기울기로 변형하여 예측하여 그 값을 이전 프레임 신호의 절대치 합으로 나눈 후 2에서 빼줌에 따라 항상 0~1사이의 값을 가지게 된다. 프레임 손실이 계속될 때 계산된 GI, GD에 따라 <그림 4>와 같이 적응적으로 이득 조절을 수행한다. 기존 G.711 PLC의 이득 조절과 유사하게 연속된 손실이 발생할수록 점점 복원된 신호의 크기를 감쇠 시키며, 기존의 방식은 <그림 2>와 같이 항상 20%의 감쇠인자를 사용하여 감쇠를 수행하지만 제안한 방법은 계산된 GI, GD에 따라 적응적으로 감쇠 비율을 결정하여 최종적으로 60ms 이상의 손실에 대해서는 묵음으로 복원한다. 여기서, 초기 이득 값의 최대 변화폭은 0.5~1.5로 제한된다.



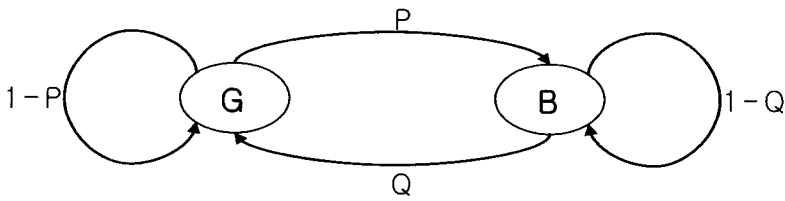
<그림 4> 적응형 이득 조절

4. 성능 평가 실험

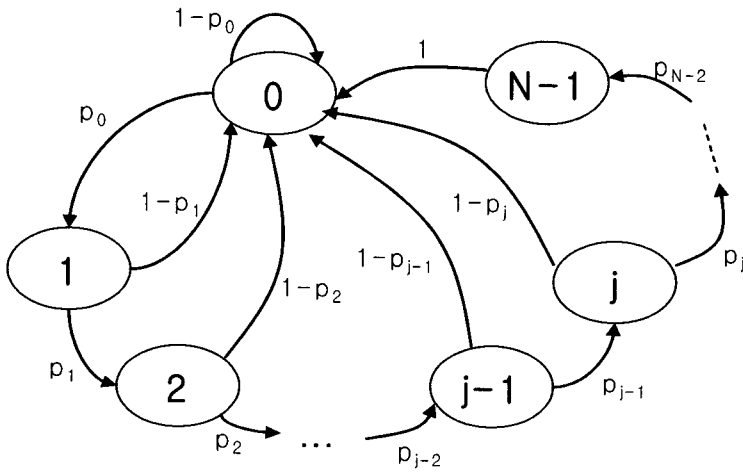
4.1. 패킷 손실 모델링

IP 네트워크에서 발생하는 손실을 모델링하기 위해 ITU-T Recommendation G.191에 포함된 STL(Software Tool Library)를 이용하였다[5]. STL은 Gilbert-Elliot 모델과 Bellcore 모델에 기반을 두어 손실형태 및 손실률 등의 파라미터 입력에 따라 랜덤 및 버스트 손실을 발생시키는 EID를 제공한다.

EID에서 Gilbert-Elliot 모델은 <그림 5>에서 보는 것과 같이 G와 B의 2개의 상태(state)를 가지는 Hidden Markov Chain으로 랜덤 손실을 모델링 하기 위해 사용된다[7]. 모델 안에는 G 또는 B 상태를 계속 유지 할 확률 $1-P$, $1-Q$ 와 이동할 확률 P , Q 가 존재하며, 각 상태에서 손실이 발생할 확률이 각각 존재한다.



<그림 5> Gilbert-Elliot 모델



<그림 6> Bellcore 모델

EID에서 Bellcore 모델은 버스트 손실의 특징을 모델링하기 위해 <그림 6>과 같이 N 개의 상태를 가지는 Markov 모델이 사용된다[8]. 어떤 상태에서 상태 0으로 이동하는 것은 손실이 발생하지 않은 경우를 나타내며 상태 $j-1$ 에서 상태 j 로 이동하는 것은 상태 j 이전의 프레임들이 손실을 가지고 있는 것을 의미하며, 각 상태에서 다른 상태와 0 상태로 이동할 확률을 가지고 있다.

본 연구에서는 다양한 패킷 손실을 가정하기 위해 1,3,5,10,15%의 랜덤 및 버스트 패킷 손실을 각각 발생시킨다. 그리고 1,2,3개의 10 ms의 음성 프레임이 한 개의 패킷을 구성하는 것으로 가정하였다. 따라서 패킷 당 프레임 수가 2, 3일 경우는 한 개의 패킷 손실이 발생하면 반드시 각각 연속적인 2, 3개의 음성 프레임 손실이 발생하는 결과를 초래한다.

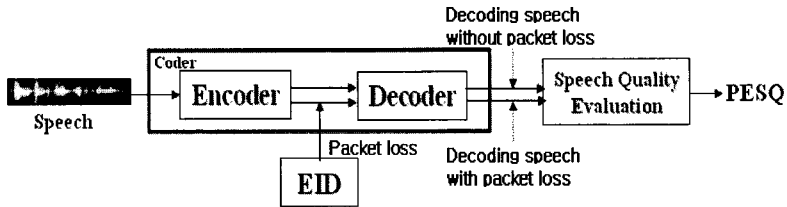
4.2. 성능 평가 방법

성능평가 방법은 ITU-T에서 제공하는 객관적인 음질평가 방법인 PESQ를 사용한다[8]. PESQ는 가장 최근에 개발된 객관적 음질 평가방법으로 기준음성과 열화

된 음성과의 비교를 통해 열화 정도를 -0.5~4.5 사이의 점수로 평가한다. 이 PESQ 값은 주관적인 평가 방법인 MOS(Mean Opinion Score)와 큰 상관도를 가지도록 개발되었다.

실험 데이터는 10초 정도의 발성 길이를 가지는 한국어 대화체 음성이 사용되었으며, 남, 여 각각 2명의 화자가 발성한 총 200개의 음성 데이터가 성능평가를 위해 사용되었다.

성능 평가는 <그림 7>과 같은 절차로 수행된다. 입력 음성 데이터에 대해 코덱에서 부호화를 수행한다. 부호화 된 데이터에 대해 EID를 통해 패킷 손실을 모델링하며, 코덱은 패킷 손실이 발생했을 때 패킷 손실 은닉 알고리즘을 수행하여 복호화를 수행한다. 음질평가를 위해 패킷 손실 없이 복호화 된 음성, 패킷 손실이 발생했을 때 기존의 패킷 손실 은닉 알고리즘을 이용하여 복호화된 음성, 패킷 손실이 발생했을 때 기존의 패킷 손실 은닉 알고리즘을 이용하여 복호화된 음성을 각각 비교하여 PESQ 값을 산출한다.



<그림 7> 실험절차

5. 성능평가 및 결과 분석

기존 G.711에 내장된 PLC 알고리즘의 성능평가 결과는 <표 1>과 같다. 표에서 1fpp(frame per packet), 2fpp, 3fpp는 각각 한 패킷이 1, 2, 3개의 프레임으로 구성된 것을 나타낸다. 결과를 보면, 패킷 손실률이 증가 할수록 음질의 저하가 크게 나타나는 것을 알 수 있으며, 랜덤 손실보다는 버스트 손실에서 음질의 저하가 큰 것을 알 수 있다. 또한, 패킷화의 영향으로 일반적으로 손실률이 작은 경우를 제외하고는 패킷 당 프레임 수가 증가 할수록 같은 손실률에 대해 음질의 저하가 큰 것을 알 수 있다.

<표 1> G.711 PLC 성능평가 결과 (PESQ 값)

손실 형태		1fpp	2fpp	3fpp
랜덤	1%	3.9825	3.9867	4.0178
	3%	3.6443	3.6302	3.6209
	5%	3.4399	3.4121	3.3896
	10%	3.1242	3.0379	2.9645
	15%	2.8934	2.7841	2.6669
버스트	1%	3.9753	3.9906	3.9833
	3%	3.5578	3.5889	3.5999
	5%	3.3167	3.2619	3.3048
	10%	2.8389	2.7583	2.7555
	15%	2.5039	2.3810	2.4071

첫 번째 제안한 방법인 유/무성음 분류에 따른 적응적 손실 양단 정보 이용을 적용했을 때의 결과는 <표 2>와 같다. <표 3>은 <표 2>의 값에서 <표 1>의 값을 빼준 것으로 성능 향상 정도를 나타낸다. 양의 값으로 커질수록 성능향상의 정도가 더 큰 것을 의미한다. 3fpp의 버스트 3%의 손실을 제외하고는 모든 손실률에서 성능이 향상된 것을 알 수 있다. 1fpp에서는 랜덤 손실에 대해 평균 0.0183, 버스트 손실에 대해 평균 0.0976 정도의 성능 향상을 보였으며, 2fpp의 경우는 랜덤 손실에 대해서는 평균 0.0902, 버스트 손실에 대해서는 평균 0.0408 정도의 성능 향상을 보였다. 3fpp의 경우는 랜덤 손실에 대해서는 평균 0.1047, 버스트 손실에 대해서는 평균 0.0172 정도의 성능 향상 정도를 보인다.

<표 2> 적응적 손실 양단 정보 이용을 적용한 PLC의 성능평가 결과 (PESQ 값)

손실 형태		1fpp	2fpp	3fpp
랜덤	1%	3.9842	4.0111	4.0563
	3%	3.65	3.6989	3.7044
	5%	3.4503	3.4999	3.4901
	10%	3.1491	3.1594	3.1059
	15%	2.9421	2.9326	2.8263
버스트	1%	4.0195	4.0227	4.0055
	3%	3.6316	3.613	3.5990
	5%	3.4065	3.2907	3.3157
	10%	2.9644	2.8110	2.7742
	15%	2.6585	2.4473	2.4423

<표 3> 적응적 손실 양단 정보 이용을 적용했을 때의 PLC 성능 향상 정도

손실 형태		1fpp	2fpp	3fpp
랜덤	1%	0.0017	0.0244	0.0385
	3%	0.0057	0.0687	0.0835
	5%	0.0104	0.0878	0.1005
	10%	0.0249	0.1215	0.1414
	15%	0.0487	0.1485	0.1594
버스트	1%	0.0442	0.0321	0.0222
	3%	0.0738	0.0241	-0.0009
	5%	0.0898	0.0288	0.0109
	10%	0.1255	0.0527	0.0187
	15%	0.1546	0.0663	0.0352

두 번째 제안한 방법인 적응적 이득 조절을 추가적으로 적용했을 때의 결과는 <표 4>와 같다. <표 5>는 <표 4>의 값에서 <표 1>의 값을 빼준 것으로 성능 향상 정도를 나타낸다. 패킷 당 프레임수가 1개일 경우는 초기 이득 값을 0.9~1.1로 제한하여 변화 시킬 때 가장 큰 성능의 향상을 보였으며, 패킷 당 프레임수가 2,3개일 경우는 각각 초기 이득 값을 0.6~1.4, 0.5~1.5로 제한하여 변화 시킬 때 가장 큰 성능의 향상을 보였다.

<표 4> 적응적 양단 정보 이용 및 적응적 이득 조절을 적용한 PLC의 성능평가 결과 (PESQ 값)

손실 형태		1fpp	2fpp	3fpp
랜덤	1%	3.9843	4.0128	4.0583
	3%	3.6514	3.7007	3.7091
	5%	3.4517	3.5031	3.4959
	10%	3.1500	3.1632	3.1127
	15%	2.9429	2.9364	2.8303
버스트	1%	4.0205	4.0255	4.0048
	3%	3.6321	3.6159	3.5972
	5%	3.4083	3.2927	3.3186
	10%	2.9659	2.8126	2.7761
	15%	2.6604	2.4476	2.4439

<표 5> 적응적 양단 정보 이용 및 적응적 이득 조절을 적용했을 때의 PLC 성능 향상 정도

손실 형태		1fpp	2fpp	3fpp
랜덤	1%	0.0018	0.0261	0.0405
	3%	0.0071	0.0705	0.0882
	5%	0.0118	0.0910	0.1063
	10%	0.0258	0.1253	0.1482
	15%	0.0495	0.1523	0.1634
버스트	1%	0.0452	0.0349	0.0215
	3%	0.0743	0.0270	-0.0027
	5%	0.0916	0.0308	0.0138
	10%	0.0127	0.0543	0.0206
	15%	0.1565	0.0666	0.0368

3fpp의 버스트 3%의 손실을 제외하고는 모든 손실률에서 성능이 향상 된 것을 알 수 있다. 1fpp에서는 랜덤 손실에 대해 평균 0.0192, 버스트 손실에 대해 평균 0.0989 정도의 성능 향상을 보였으며, 2fpp의 경우는 랜덤 손실에 대해서는 평균 0.0930, 버스트 손실에 대해서는 평균 0.0427 정도의 성능 향상을 보였다. 3fpp의 경우는 랜덤 손실에 대해서는 평균 0.1093, 버스트 손실에 대해서는 평균 0.018 정도의 성능 향상 정도를 보인다.

<표 6> 적응적 이득 조절을 추가적으로 적용했을 때 추가적인 PLC 성능 향상 정도

손실 형태		1fpp	2fpp	3fpp
랜덤	1%	0.0001	0.0017	0.0020
	3%	0.0014	0.0018	0.0047
	5%	0.0014	0.0032	0.0058
	10%	0.0009	0.0038	0.0068
	15%	0.0008	0.0038	0.0040
버스트	1%	0.0010	0.0028	-0.0070
	3%	0.0005	0.0029	-0.0018
	5%	0.0018	0.0020	0.0029
	10%	0.0015	0.0016	0.0019
	15%	0.0019	0.0003	0.0016

<표 6>은 <표 4>의 값에서 <표 2>의 값을 빼준 것으로 유/무성음 분류에 따른

적응적 손실 양단 정보 이용을 적용하고 적응형 이득 조절을 적용했을 때의 추가적인 성능향상 정도를 나타낸다. 표에서 보는 것과 같이 추가적인 성능 향상을 보임을 알 수 있고, 1fpp의 랜덤 손실에서는 평균 0.0009, 버스트 손실에 대해서는 평균 0.0013 정도의 성능 향상을 보였다. 2fpp의 경우는 랜덤 손실에 대해 평균 0.0029, 버스트 손실에 대해 평균 0.0031의 추가적인 성능 향상을 보였으며, 3fpp에서는 랜덤 손실에 대해 평균 0.0047, 버스트 손실에 대해 평균 0.0041의 추가적인 성능 향상을 보였다.

전체적인 성능 향상 정도를 분석해 보면 1fpp에서는 랜덤보다는 버스트에서 성능향상 정도가 큰 것을 알 수 있고, 손실률이 증가할수록 성능향상 정도가 커진다. 2,3fpp의 경우에서도 역시 손실률이 증가할수록 성능향상의 정도가 커지는 것을 알 수 있다. 따라서 제안한 PLC 알고리즘은 기존의 PLC 알고리즘보다 버스트 손실에서 더욱 강인한 특징을 보임을 알 수 있다. 그런데 2,3 fpp에서는 버스트 손실에서 보다는 랜덤 손실에서 성능향상 정도가 큰 것을 알 수 있는데, 이는 패킷화의 영향으로 분석된다. 즉, 제안된 알고리즘이 버스트 손실에 강인한 특징을 보이는데, 2,3fpp의 랜덤 손실 역시 부분적으로는 버스트 패킷 손실로 볼 수 있기 때문이다. 왜냐하면 2,3fpp에서는 한번에 항상 2개 혹은 3개의 손실이 연속적으로 발생하기 때문이다. 반대로 2,3fpp의 버스트 손실은 패킷화의 영향으로 상당히 긴 프레임 손실이 발생함에 따라 손실을 묵음 처리해 줌에 따라 성능 향상 정도가 크지 않음을 알 수 있다.

6. 결 론

본 논문에서는 G.711의 패킷 손실 은닉 알고리즘의 성능향상을 위해 음성의 특징을 이용한 유/무성음 분류에 따른 적응적 손실 양단 정보 이용, 적응형 이득 조절 방법을 이용한 패킷 손실 보상 방법을 제안하였다. 제안된 방법은 G.711 Appendix I 패킷 손실 은닉 알고리즘과 객관적인 음질 평가 방법인 PESQ 실험을 통해 비교되었다. 실제 통신 환경에서 발생할 수 있는 다양한 패킷 손실 환경을 가정하기 위해 각각 1, 3, 5, 10, 15%의 손실률을 가지는 랜덤, 버스트 손실을 발생시켰으며 패킷 당 프레임수를 1, 2, 3개로 가정하여 실험을 수행 하였다.

성능비교 결과 본 연구에서 제안한 알고리즘을 사용했을 때 기존의 방법보다 패킷 손실 은닉 알고리즘의 성능이 향상되는 것을 확인하였다. 또한 제안한 패킷 손실 은닉 알고리즘은 버스트 패킷 손실에 더욱 강인한 특징을 보임을 확인하였다. 따라서 제안된 패킷 손실 은닉 알고리즘을 내장한 코덱을 VoIP 서비스에 이용하면 패킷 손실이 발생했을 때 보다 높은 통화 품질을 유지하면서 서비스를 제공할 수 있다. 따라서 VoIP 서비스의 가장 큰 이슈들 중에 하나인 QoS 보장 문제에

있어서, 패킷 손실에 따른 음질 저하 문제의 해결책의 한가지 방법으로 이용 할 수 있다.

본 논문에서 유/무성음 분류를 위해 CART를 이용한 결정트리 방식을 사용하였는데, 이 방식은 훈련데이터에 종속적인 결정트리가 생성되는 특징이 있으므로 실제 환경과 유사한 훈련데이터를 이용하여 훈련하는 것이 중요하다. 따라서, 훈련 데이터로써 실제 VoIP 환경과 유사한 데이터를 사용하는 것이 중요하며, 유/무성음 분류를 위해 결정트리 방식이 아닌 다른 방법을 적용할 수 있을 것이다. 그리고 앞으로의 연구로써 유/무성음 분류 오류가 PLC 성능에 미치는 영향에 대한 계속적인 연구가 필요하다고 생각된다.

참 고 문 헌

- [1] ITU-T Recommendation, Appendix I: A high quality low-complexity algorithm for packet loss concealment with G.711, *ITU-T Recommendation G.711*, Sep. 1999.
- [2] ITU-T Recommendation, Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s, *ITU-T Recommendation G.723.1*, Mar. 1996.
- [3] ITU-T Recommendation, Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited-linear-prediction (CS-ACELP), *ITU-T Recommendation G.729 Annex A*, Nov. 1996.
- [4] ITU-T Recommendation, Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs, *ITU-T Recommendation P.862*, Feb. 2001.
- [5] ITU-T Recommendation, Software tools for speech and audio coding standardization, *ITU-T Recommendation G.191*, Nov. 2000.
- [6] R. O. Duda, P. E. Hart, et al., *Pattern Classification*, JOHN WILEY & SON, 2001.
- [7] M. Mushkin, I. Bar-David, "Capacity and coding for the Gilbert-Elliot channels", *IEEE Transactions on Information Theory*, Vol. 35, No. 6, pp. 1277-1290, Nov. 1989.
- [8] V. K. Varma, "Testing speech coders for usage in wireless communications systems," *Proc. of IEEE Workshop on Speech Coding for Telecommunication*, pp. 94-94, Oct. 1993.

접수일자: 2006년 2월 20일

게재결정: 2006년 3월 16일

▶ 한승호(Seung Ho Han)

주소: 305-732 대전시 유성구 문지동 103-6 한국정보통신대학교

소속: 한국정보통신대학교(ICU) 음성/음향 정보 연구실

전화: 042) 866-6204

E-mail: space0128@icu.ac.kr

▶ 김진술(Jinsul Kim)

주소: 305-700 대전시 유성구 가정동 161 한국전자통신연구원

소속: 한국전자통신연구원(ETRI) 광대역통합망연구단 BcN 연동기술팀

전화: 042) 860-6587

E-mail: jsetri@etri.re.kr

▶ 이현우(Hyun-Woo Lee)

주소: 305-700 대전시 유성구 가정동 161 한국전자통신연구원

소속: 한국전자통신연구원(ETRI) 광대역통합망연구단 BcN 연동기술팀

전화: 042) 860-6526

E-mail: hwlee@etri.re.kr

▶ 류원(Won Ryu)

주소: 305-700 대전시 유성구 가정동 161 한국전자통신연구원

소속: 한국전자통신연구원(ETRI) 광대역통합망연구단 BcN 연동기술팀

전화: 042) 860-6153

E-mail: wlyu@etri.re.kr

▶ 한민수(Minsoo Hahn) : 교신저자

주소: 305-732 대전시 유성구 문지동 103-6 한국정보통신대학교

소속: 한국정보통신대학교(ICU) 음성/음향 정보 연구실

전화: 042) 866-6123

E-mail: mshahn@icu.ac.kr