

적응적 신호 크기 예측을 이용한 G.711 패킷 손실 은닉 알고리즘의 성능향상

Performance Improvement of Packet Loss Concealment Algorithm in G.711 Using Adaptive Signal Scale Estimation

김태하, 이인성[†]

(Tae-Ha Kim and In-Sung Lee[†])

충북대학교 전파통신공학과

(Received June 15, 2015; accepted July 30, 2015)

초 록: 본 논문에서는 G.711 패킷 손실 은닉 알고리즘의 성능향상을 위해 적응적 신호 크기 예측을 사용하는 패킷 손실 알고리즘을 제안한다. 기존의 방법은 연속 손실이 발생하였을 때 20%의 감쇠인자를 가지고 이득조절을 수행하였다. 그러나 이 방법은 신호의 변화를 고려하지 않기 때문에 신호가 왜곡되는 경우가 발생한다. 따라서 Least Mean Square(LMS) 예측기를 사용하여 이전과 이후 프레임의 정보를 통한 적응적 신호 크기 예측으로 이득을 조절하는 것을 제안한다. 제안된 알고리즘의 성능 평가는 Perceptual Evaluation of Speech Quality(PESQ) 평가를 통하여 나타내었다. **핵심어:** 패킷 손실 은닉, LMS 예측기, G.711, PESQ, 이득조정

ABSTRACT: In this paper, we propose Packet Loss Concealment (PLC) method using adaptive signal scale estimation for performance improvement of G.711 PLC. The conventional method controls a gain using 20% attenuation factor when continuous loss occurs. However, this method lead to deterioration because that don't consider the change of signal. So, we propose gain control by adaptive signal scale estimation through before and after frame information using Least Mean Square (LMS) predictor. Performance evaluation of proposed algorithm is presented through Perceptual Evaluation of Speech Quality (PESQ) evaluation.

Keywords: PLC, LMS, Predictor, G.711, PESQ, Gain control

PACS numbers: 43.72.Kb, 43.60.Dh

1. 서 론

통신 서비스사업의 신 성장 서비스로서 VoIP(Voice over IP)는 최근 사용고객 수에 있어서 상당한 증가를 보여주고 있다. 음성 이외에도 다양한 서비스가 제공되고 있다. 따라서 이러한 VoIP 서비스의 확대 공급에 따라 서비스의 품질관리가 중요한 이슈로 부각되고 있다.^[1]

IP 네트워크를 이용하여 패킷을 기반으로 음성통신을 수행하는 VoIP 서비스의 경우 IP망을 기반으로

음성을 전송하기 때문에 지연, 지터, 패킷 손실 등으로 인하여 QoS(Quality of Service)를 보장받지 못하기 때문에 품질 저하를 유발하게 된다.^[2]

네트워크 망을 사용하는 표준 음성 코덱(G.711, G.723.1, G.729 등)들은 자체적으로 손실된 패킷을 복구하여 손실이 나지 않은 것처럼 보이게 하는 패킷 손실 은닉(Packet Loss Concealment, PLC) 알고리즘을 사용하고 있다.^[4]

패킷 손실 은닉 알고리즘의 경우 두 가지 종류가 있다. 먼저 수신자 측에서 손실을 알리기 위해 데이터에 미리 정보를 포함하여 전송 후 수신자 측에서 그 정보를 받아 손실된 패킷을 복구하는 송신자와 수신자 기반의 방법이 있고 수신자 측에서 손실을

[†]Corresponding author: In-Sung Lee (inslee@chungbuk.ac.kr)
Department of Radio Communication Eng Chungbuk National University, Gaesung-dong, Seowon-gu, Cheongju 28644, Republic of Korea
(Tel: 82-43-261-3223)

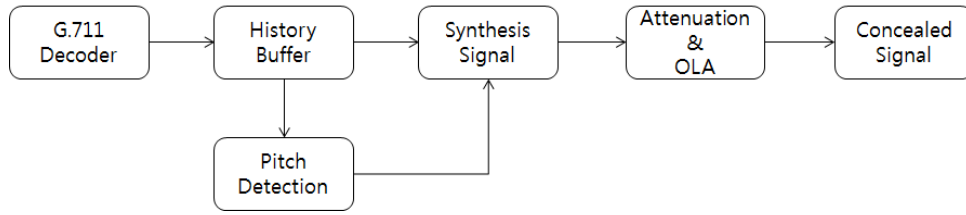


Fig. 1. G.711 PLC algorithm.

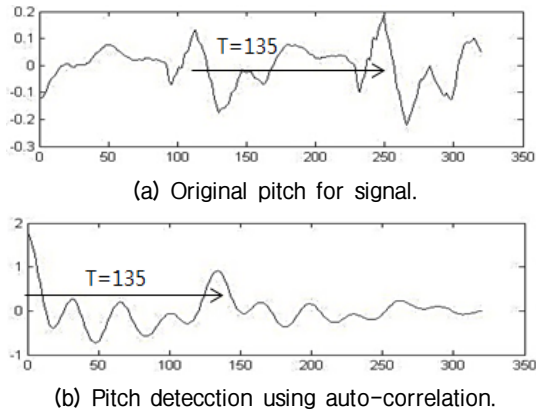


Fig. 2. Pitch detection.

확인 후 자체적인 방법으로 손실을 복구하는 수신자 기반의 방법^[3]이 있다.

표준 음성 코덱 중 하나인 G.711 코덱의 경우 후자인 수신자 기반의 방법을 사용한다. PCM(Pulse Code Modulation) 기반의 코덱이기 때문에 음성 데이터를 직접 이용하여 PLC를 수행하게 된다. 손실이 발생한 경우 히스토리 버퍼에 저장되어있는 음성 신호로부터 피치 구간을 계산한 후 계산된 피치 구간을 반복, 삽입 하여 음성을 복원하게 된다. 최종적으로 계산된 피치의 1/4 지점을 OLA(OverLap Add)를 이용하여 음성이 부드럽게 될 수 있도록 만들어 준다. 연속 손실이 발생하였을 경우에는 20%의 감쇠인자를 이용하여 이득조절을 수행하게 되고 60 ms 이상의 손실이 발생하게 되면 묵음으로 처리하게 된다.

이러한 방법은 신호의 크기 변화를 고려하지 않고 일방적으로 감쇠만 시키게 되므로 변화되는 신호에 대해서 신호의 왜곡을 가져올 수가 있다.

기존의 G.711 PLC 알고리즘의 성능을 향상시키기 위한 방법으로는 LP(Linear Prediction)계수와 손실 이후 프레임을 이용하여 각각 단일 프레임과 연속 손실 프레임을 처리하는 방법,^[10] CART(Classification

And Regression Tree) 알고리즘을 사용하여 유/무성음을 분리한 후 손실 양단의 정보를 이용한 방법^[2] 등이 제시 되어 있다.

본 논문에서는 손실된 프레임을 기준으로 양쪽 프레임의 최대값을 활용하여 신호의 크기 변화를 감지하여 적응적으로 이득조절을 수행하는 방법을 제안하였다. 이 과정에서 양쪽 프레임 중 하나라도 손실이 발생하였다면 예측기를 이용하여 예측된 최대값을 활용하여 이득조절을 수행하였고, 그렇지 않은 경우에는 손실되지 않은 양쪽 프레임의 정보를 활용하여 이득조절을 수행하였다. 기존의 G.711 PLC 알고리즘과 제안된 알고리즘의 성능평가는 ITU-T에서 제공하는 PESQ(Perceptual Evaluation of Speech Quality)로 각각의 결과 값을 측정하였다.

II. G.711 패킷 손실 은닉 알고리즘

G.711 코덱의 경우 PCM 기반의 코덱이기 때문에 직접 음성 데이터를 가지고 PLC 알고리즘을 수행하게 된다. Fig. 1은 G.711 PLC 알고리즘을 나타낸다. 히스토리 버퍼에 저장되어있는 48.75 ms 이전의 음성 신호로부터 피치 구간을 계산하게 되는데 피치의 경우 Eq.(1) 처럼 자기상관함수^[6]를 사용하여 구할 수 있다.

$$R[k] = \sum_{m=0}^{L-k} x[m] \cdot x[m-k], \quad k=0, \dots, 300. \quad (1)$$

$$\hat{S}_k^{recom} = \hat{S}_{k-1}(L - P_k + n), \quad n=0, \dots, P. \quad (2)$$

Eq.(1)을 이용하여 구한 피치의 예를 Fig. 2에서 보여주고 있다. 다음과 같이 구해진 피치를 이용하여 Eq.(2)를 이용하여 프레임 길이만큼 반복, 삽입 하여

합성한 후에 최종적으로 계산된 피치의 1/4지점을 Eq.(1)을 이용하여 구한 피치의 예를 Fig. 2에서 보여 주고 있다. 다음과 같이 구해진 피치를 이용하여 Eq.(2)를 이용하여 프레임 길이만큼 반복, 삽입하여 합성한 후에 최종적으로 계산된 피치의 1/4지점을 OLA하여 손실된 음성을 복원하게 된다. Eq.(2)를 이용하여 신호를 복원한 신호를 Fig. 3에서 보여주고 있다.

연속 손실이 일어난 경우에는 Fig. 4의 PLC 알고리즘이 적용된다. 첫 번째 손실의 경우 기존의 알고리즘과 동일하게 적용이 되고, 두 번째 손실부터는 연

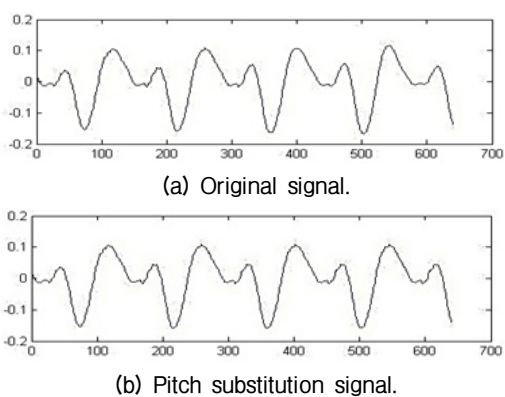


Fig. 3. Pitch substitution.

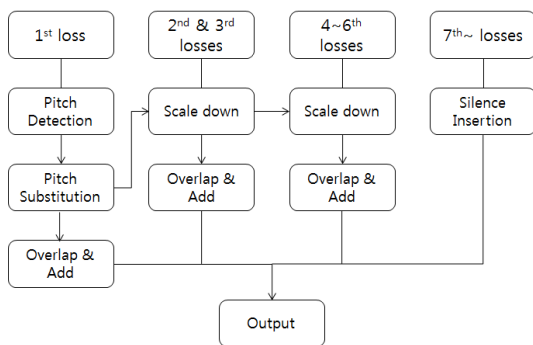


Fig. 4. G.711 PLC algorithm for continuous packet loss.

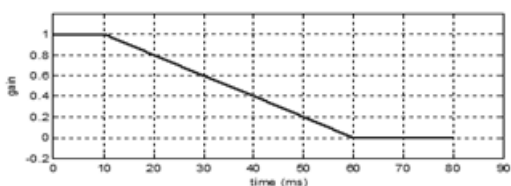


Fig. 5. G.711 PLC gain control.

속손실로 간주하여 감쇠 인자를 적용하게 된다. 두 번째, 세 번째 연속손실의 경우 첫 번째 손실이 발생하였을 때 구한 피치를 이용하여 피치의 삽입, 반복

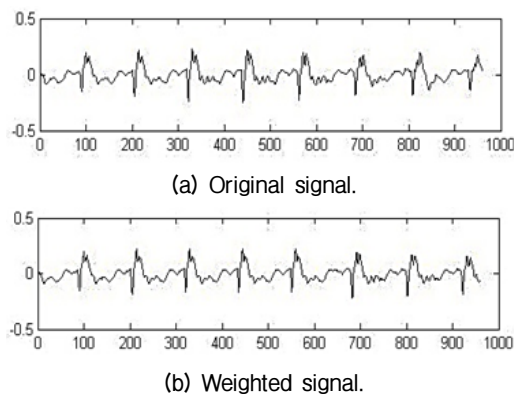


Fig. 6. Weighted signal using attenuation factor for 2nd loss.

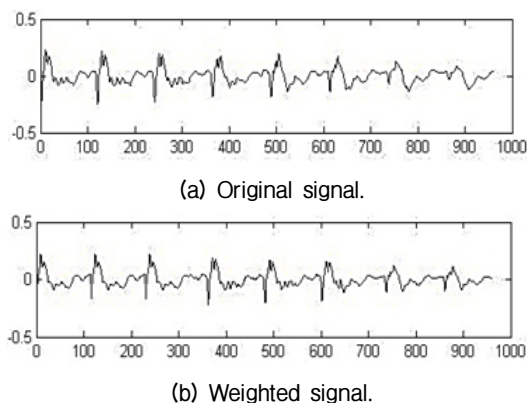


Fig. 7. Weighted signal using attenuation factor for 3rd loss.

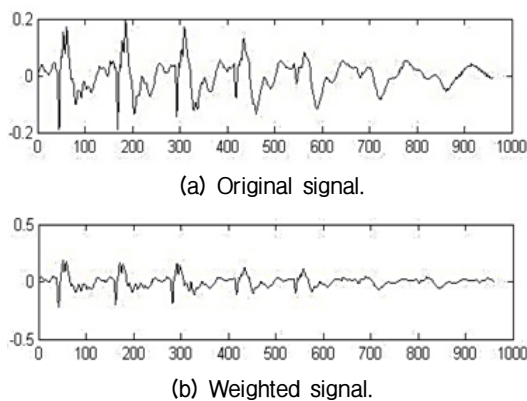


Fig. 8. Weighted signal using attenuation factor for 4th loss.

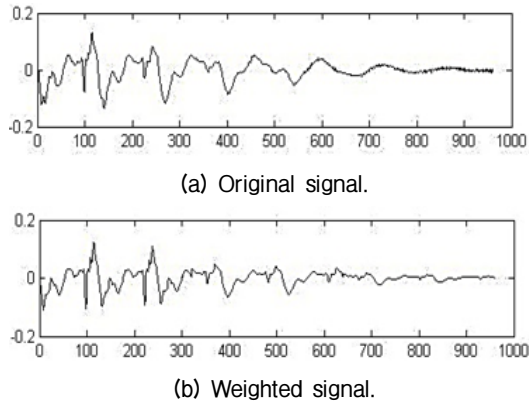


Fig. 9. Weighted signal using attenuation factor for 5th loss.

으로 생성된 신호에 감쇠 인자를 적용하여 이득을 조절한 후 신호를 복원하게 된다. 네 번째, 다섯 번째 연속 손실의 경우 세 번째 손실에서 감쇠된 신호에 다시 한 번 감쇠 인자를 적용하여 이득을 조절한 후에 신호를 복원하게 되고 일곱 번째 손실부터는 묵음으로 처리하게 된다. 감쇠 인자의 경우 Fig. 5에서 보이듯이 연속적으로 손실이 발생할 때 마다 20%씩 줄어들면서 이득 조절이 수행된다. Figs. 6, 7, 8, 9에서는 연속 손실이 발생한 경우 감쇠 인자로 신호의 이득을 조절한 신호를 보여주고 있다. 연속 손실이 발생할 때마다 감쇠 인자가 적용되어 신호가 점점 감쇠하는 것을 볼 수가 있다.

III. 적응적 신호 크기 예측을 통한 이득조절

3.1 LMS(Least Mean Square) 예측기를 통한 최대값 예측

LMS 예측기는 적응 필터의 한 종류인 LMS 알고리즘의 예측 기능을 이용한 것이다. 이전 프레임의 최대값, 스텝크기, 예측오차를 이용하여 필터 계수를 갱신하고 갱신된 필터계수와 이전프레임의 최대값 곱으로 현재 프레임의 최대값을 예측할 수 있다.^[7] Eq.(3)을 이용하여 필터계수가 갱신되고 Eq.(4)를 이용하여 예측 오차를 구할 수 있으며 최종적으로 Eq.(5)를 이용하여 손실된 프레임의 최대값을 예측할 수가 있다.

$$H_{k+1} = H_k + \mu \cdot E_k \cdot M_{k-1}, \quad (3)$$

$$E_k = M_k - \hat{M}_{k-1}, \quad (4)$$

$$\hat{M}_k = H_k \cdot M_{k-1}, \quad (5)$$

여기서 μ 는 step size, k 는 프레임 인덱스, E_k 는 예측 오차, M_{k-1} 은 이전 프레임의 최대값을 나타낸다. 다음 식들을 사용하여 얻은 최대값 예측치를 Fig. 10에서 보여 주고 있다.

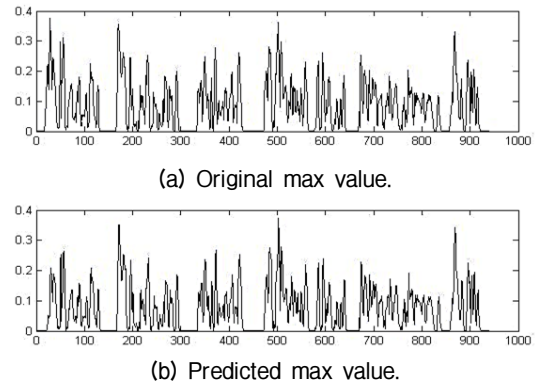


Fig. 10. Max value prediction using LMS predictor.

3.2 적응적 신호 크기 예측을 통한 이득조절

음성의 경우 시간에 대해서 천천히 변화하는 특성을 가지고 있다. 따라서 손실된 음성과 가장 가까운 구간의 음성이 손실된 음성과 가장 유사한 특성을 가지게 된다. 따라서 손실된 프레임을 중심으로 이전 신호에 비해 이후 신호가 커졌다면 그 손실된 프레임도 커지는 특성을 가지게 될 것이고 이전신호에 비해 이후 신호가 작아졌다면 그 손실된 프레임도 작아지는 특성을 가지게 될 것이다. 따라서 손실된 프레임을 기준으로 양쪽 프레임의 최대값 정보를 활용하여 이득을 조절한다.

Fig. 11에서는 증가하는 신호에 적용하는 W_i 와 감소하는 신호에 적용하는 W_d 를 보여준다. 이전 최대값에 비해 이후 최대값이 큰 경우에는 W_i 를 적용하여 이득을 조절하고 그 반대의 경우에는 W_d 를 적용하여 이득을 조절하게 된다. W_i 와 W_d 의 경우 프레임의 최대값을 활용하여 Eq.(6)으로 구하게 된다.

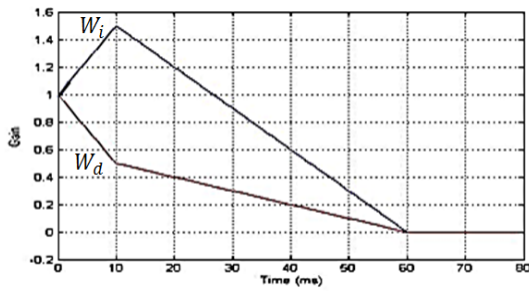


Fig. 11. Gain control using adaptive signal scale estimation.

$$W = \begin{cases} W_i = \frac{\hat{M}_{k+1}}{\hat{M}_{k-1}} & (if \hat{M}_{k+1} > \hat{M}_{k-1}), \\ W_d = \frac{\hat{M}_{k+1}}{\hat{M}_{k-1}} & (if \hat{M}_{k+1} < \hat{M}_{k-1}). \end{cases} \quad (6)$$

손실된 프레임을 기준으로 양쪽 프레임의 손실이 없으면 양쪽프레임의 최대값을 Eq.(6)에 적용하고 양쪽 프레임에 손실이 하나라도 생겼다면 LMS 예측기를 사용하여 구한 예측된 최대값을 Eq.(6)에 적용하게 된다.

Figs. 12와 13은 각각 증가하는 신호와 감소하는 신호에 대해서 Fig. 11의 이득조절을 적용한 것이다. Fig. 12의 경우 W_i 를 적용하여 이득조절을 한 것이고 Fig. 13은 W_d 를 적용하여 이득조절 한 것이다. 증가하는 신호에 대해서 증가하는 형태로 신호가 복원되었고 감소하는 신호에 대해서 감소하는 형태로 신호가 복원된 모습을 볼 수가 있다.

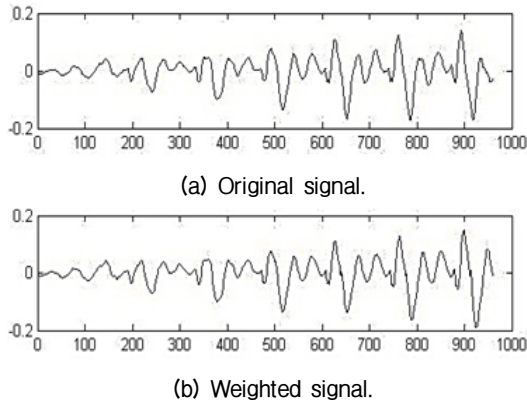


Fig. 12. Weighted signal using proposed method for increasing signal.

IV. 성능평가

본 논문에서 제안하는 알고리즘을 실험하기 위해서 ITU-T에서 제공하는 객관적인 음질 평가 방법인 Perceptual Evaluation of Speech Quality(PESQ) 방법을 사용하였다.^[8] 또한 IP 네트워크 상에서 발생하는 손실을 모델링 하기 위하여 Gilbert-Elliot Model을 사용하였다.^[9] Gilbert-Elliot Model의 경우 Fig. 14에서 보이는 것과 같이 Markov Chain으로 표현이 되어있다. 평균 손실률과 연속 손실 확률은 Eqs.(7)과 (8)의 식으로 계산할 수 있다.

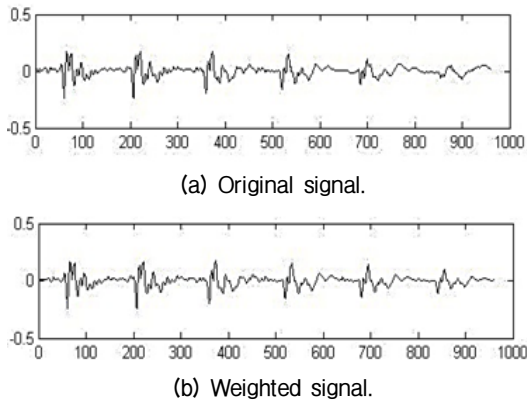


Fig. 13. Weighted signal using proposed method for decreasing signal.

$$L_{avg} = \frac{P1}{P1 + P2}. \quad (7)$$

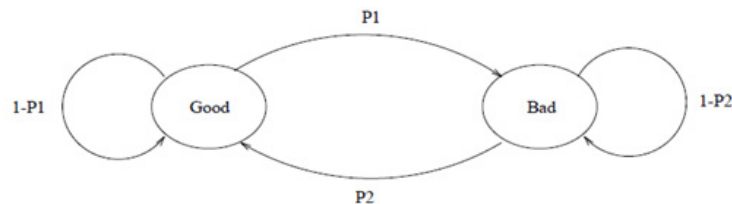


Fig. 14. Gilbert-Elliot model.

$$L_{n-continuous} = (1 - P2) \cdot P2^{n-1}. \quad (8)$$

평균 손실률의 경우 다양한 손실률에서 제안한 알고리즘을 평가하기 위해 5%, 10%, 20%, 30%의 손실률을 적용하였고 연속 손실률의 경우 한 번 연속 손실 될 확률을 50%로 하여 적용하였다. 손실률에 따른 P1과 P2는 Table 1에 표기하였다.

실험에 사용한 입력신호는 Table 2에 정의하였다. 이 신호는 ITU-T에서 제공하는 G.711 코덱^[5]을 사용하여 인코딩, 디코딩 과정을 적용하였다.

Table 3과 Table 4에서는 제안할 알고리즘을 적용한 신호의 PESQ 값을 보여주고 있다. G.711 코덱을

Table 1. Probability of Error Model.

Average Loss Rate	P1	P2
5 %	0.0263	0.5
10 %	0.05555	0.5
20 %	0.125	0.5
30 %	0.21459	0.5
40 %	0.3333	0.5

Table 2. Test Speech Signal.

File Name	Sampling Rate	Total Length	Frame Length
Korean_male1.wav	16 kHz	18.760 s	20 ms
Korean_male2.wav	16 kHz	14.120 s	20 ms
Korean_female1.wav	16 kHz	18.980 s	20 ms
Korean_female2.wav	16 kHz	15.020 s	20 ms

Table 3. PESQ Result (Korean Male).

Loss Rate	G.711 PLC	Proposed PLC
5 %	3.4555	3.5222
10 %	3.1472	3.2155
20 %	2.6787	2.7502
30 %	2.4401	2.5247
40 %	2.3626	2.4782

Table 4. PESQ Result (Korean Female).

Loss Rate	G.711 PLC	Proposed PLC
5 %	3.2517	3.2961
10 %	2.7587	2.8064
20 %	2.3952	2.4561
30 %	2.0138	2.1013
40 %	1.5487	1.6510

거쳐서 나온 한국남성음성과 한국여성음성의 PESQ 값은 각각 4.2256, 4.0771였으며 Error Model을 적용하여 손실을 만든 후 기존의 알고리즘과 제안한 알고리즘을 적용하였다.

일방적인 감쇠를 적용하여 이득을 조절한 기존의 방법 보다 적응적인 신호 크기 예측을 통해 이득을 조절해주었기 때문에 기존 알고리즘보다 왜곡이 줄어들게 되고 그에 따라 향상된 결과를 얻을 수 있었다.

V. 결 론

본 논문에서는 G.711 PLC 알고리즘의 성능향상을 위하여 이득 조절시 나타나는 단점을 LMS 예측기와 적응적인 신호 크기 예측에 따른 이득조절로서 보완하였다. 손실된 프레임의 기점으로 양쪽 프레임의 최대값 정보를 활용함으로써 손실된 신호의 크기변화를 예측할 수 있었고 그에 따라 적절한 이득조절이 수행됨으로써 기존 알고리즘보다 나아진 성능을 기대하였다.

결과적으로 PESQ 성능평가를 통하여 기존의 알고리즘의 성능보다 제안된 알고리즘의 성능이 남성음성의 경우 최대 0.1156만큼 향상되었고 여성음성의 경우 최대 0.1023만큼 향상된 성능을 확인할 수 있었다. 또한 손실률이 증가할수록 손실이 발생한 부분을 일방적인 감쇠가 아닌 적응적인 신호 크기 예측을 통한 이득조절로서 기존의 알고리즘의 성능보다 제안한 알고리즘의 성능이 향상되는 것을 볼 수가 있었다.

감사의 글

이 논문은 2014년도 충북대학교 학술연구지원사업의 연구비 지원에 의하여 연구되었음.

References

1. B. Y. Chang, D. W. Seo, and B. J. Park, "Study on VoIP Service Quality Management" (in Korean), J. the Institute of Webcasting, Internet and Telecommunication **11**, 245-252 (2011).
2. S. H. Han, J. S. Kim, H. W. Lee, W. Ryu, and M. S. Hahn,

- “Performance Improvement of Packet Loss Concealment Algorithm in G.711 Using Speech Characteristics” (in Korean), *Phonetics and Speech Sciences* **57**, 175-198 (2006).
3. V. P. Bhute and U. N. Shrawankar, “Speech Packet Concealment Techniques Based on Time-Scale Modification for VoIP;” in *IEEE International Conference on Computer Science and Information Technology (ICCSIT)*, 825-828 (2008).
 4. ITU-T Recommendation, *Appendix I: A high quality low-complexity algorithm for packet loss concealment with G.711*, 1999.
 5. ITU-T Recommendation, *Amendment 2: New Appendix III - Audio quality enhancement toolbox*, 2009.
 6. X. Mei, J. Pan, and S. Sun, “Efficient algorithms for speech pitch estimation,” in *IEEE International Symposium on Intelligent Multimedia, Video and Speech Processing (ISIMP)*, 421-424 (2001).
 7. Zeidler and J.R., “Performance analysis of LMS adaptive prediction filter,” *Proceedings of the IEEE*. **78**, 1781-1806 (1990).
 8. 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*, 2001.
 9. M. Mushkin and I. Bar-david, “Capacity and coding for the Gilbert-Elliot channels,” *IEEE transactions on Information Theory* **35**, 1277-1290 (1989).
 10. J. H. Kim and M. S. Han “Improvement of Packet Loss Concealment Algorithm by Utilizing Next Good Frame info” (in Korean), *Phonetics and Speech Sciences* **43**, 245-252 (2002).

저자 약력

▶ 김 태 하 (Taeha Kim)



2013년: 충북대학교 정보통신공학과 학사
2014년 ~ 현재: 충북대학교 전파통신공학과 석사
<주관심 분야> 음성/오디오 신호처리

▶ 이 인 성 (Insung Lee)



1983년: 연세대학교 전자공학과 학사
1985년: 연세대학교 전자공학과 석사
1992년: Texas A&M University 전기공학과 박사
1993년 ~ 1995년: 한국전자 통신연구원 이동통신 기술연구단 선임연구원
1995년 ~ 현재: 충북대학교 전파통신공학과 교수
<주관심 분야> 음성/영상 신호압축, 이동통신, 적응필터