

인터넷상의 실시간 음성 전송을 위한 에러 복원 알고리즘의 연구

신현숙* · 최연성**

군산대학교

Study of Error Reconstruction Algorithm for Real-time Voice for Transmissions over the Internet

Hyun-suk Shin* · Sung-yeon Choi**

*Kunsan National Univ. School of Electronic & Informtion Engineering

E-mail : purple@silkbox.net

요 약

인터넷상의 실시간 음성 전송 시에 발생하는 손실을 은닉하기 위한 다수의 알고리즘들이 제안되고 있다. 이 알고리즘들의 주 목적은 적은 대역폭을 사용하여 손실을 복원하고 복원 후 좋은 음질을 보장하는 것이다. 손실 은닉 알고리즘들은 receiver based와 sender- / receiver-based로 나뉘어진다. 본 논문에서 sender- 와 receiver-based 복원 알고리즘을 CELP를 사용하는 저 비트율 코덱에 적용하고자 한다.

ABSTRACT

In this paper, a large number of algorithm have been proposed for error concealment and reconstruction real-time voice transmission for over the internet. The main purpose of this algorithm perform error reconstruction using low bandwidth and then guarantee good voice quality. Error concealment algorithm can be classified into receiver-based and sender- and receiver-based. In this paper, we apply the sender - and receiver-based reconstruction algorithm to low bit rate codec using CELP.

1. 서 론

현재의 인터넷은 SONET 이나 ATM에 의해 수십 기가급 정도의 전송능력 밖에 가지고 있지 않기 때문에 폭증하는 트래픽을 수용하는데 한계가 있다. 또한 인터넷 프로토콜은 트래픽 엔지니어링 기능이 매우 취약하여 적절한 서비스 품질의 보장이 어렵다. 국내 인터넷의 경우, RTT(Round Trip Time)가 76~1,250ms 이고, 패킷 손실이 4~68%에 이르고 있어 서비스 품질이 제대로 지원되지 못하고 있다. 이러한 인터넷 환경에서 실시간 음성 전송시의 에러 은닉이 필요하다[1].

에러 은닉을 위한 알고리즘은 receiver based와 sender와 receiver based의 두 종류로 구분할 수 있다. receiver-based 복원 알고리즘들에서 손실된 패킷들에 묵음 또는 백색잡음을 삽입하거나[2], 가장 최근 수신된 패킷들을 반복하거나 pattern matching의 임의의 형태로 만든 후에 손실된 패킷들에 이전에 수신된 패킷들을 대신함[3]으로써 다시 만들어진다. 이러한 방법들은 손실이 가끔 발생할 때 그리고 프레임 크기가 작을 때[4] 잘 동작한다. sender- /receiver- based 복원 알고리즘들은 더욱 복잡하지만 훨씬 효과적이다. 일반적인 방법은 수신 측에서 잃어버린 데이터를 더 잘 복원할 수 있도록 하기 위해 송신 측에서 입력 스트림을 먼저 처리한다. 입력 데이터를 처리하는 다른 방법들을 기본으로 이들 알고리즘들은 여분

의 제어를 추가하는 것들과 그렇지 않은 것들로 나눌 수 있다. 여분을 추가하는 송신 측의 방법들이 있다. 이들은 복잡한 패킷을 송신하거나 [5] 현재 패킷과 함께 이전 패킷들을 송신[4]하거나 forward error correction(FEC)을 사용하여 음성 패킷들의 에러 징정 비트들을 함께 송신한다 ([6],[7]). 이들 방법들은 추가 대역폭을 요구하거나 end-to-end 지연을 요구한다. 데이터를 전송하는데 추가적인 대역폭을 요구하지 않는 방법들이 있다. 이 방법은 홀수 샘플과 짝수 샘플을 각각 그룹 짓는다. 이 두 패킷을 interleaving pair라고 한다. 이들 패킷은 서로 관여하지 않고 전송된다. 수신 측은 손실된 샘플을 복원하기 위해 다른 패킷의 샘플들의 평균을 이용해 쉽게 복원할 수 있다. 이 방법은 추가 대역폭을 요구하지 않고 손실된 패킷을 빠르게 복원 할 수 있지만 신호가 빨리 바뀌거나 신호가 모두 관련이 없다면 복원에 실패할 수도 있다.

본 논문에서는 빠르게 변하는 신호에 적응할 수 있도록 하기 위해 2가지 방법을 G.723에 적용하여 성능 분석을 하였다. 첫 번째는 수신 측에서 averaging을 사용한 복원 후 Adaptive filter를 적용한다. 두 번째는 신호를 전송하기 전 Adaptive filter를 이용하여 송신 측에서 데이터를 변형한다. II장에서 기존의 복원 방법에 대해 알아보고, III장에서 Adaptive-Filter-Based 복원 방법과 transformation-based 복원 방법에 대해 자세히 알아본다. IV장에서는 테스트 결과를 보인다. 마지막으로 결론과 향후 연구방향을 제시한다.

II. 본 론

이 장에서는 interleaving과 평균 복원을 기본으로 하는 Adaptive-Filter-Based 방법을 소개한다. interleaving 기법은 간단하게 송신 측에서 먼저 음성 스트림을 짝수 샘플과 홀수 샘플의 두개의 서브 스트림으로 나누는 방식이다. 이때 송신 측은 이들 서브 스트림을 패킷으로 그룹짓고 인터넷을 통해 그들을 전송하는 짝수와 홀수 샘플에 대응하는 두개의 패킷은 interleaving pair라고 한다. 손실이 없을 경우 수신 측은 두개의 서브 스트림을 수신하고 패킷을 원 신호로 de-interleave한다. interleaving pair에서 한 개의 패킷이 손실되었을 경우 수신 측은 수신된 두 개의 인접 샘플들의 평균으로 손실된 패킷을 복원한다. 각 패킷의 크기가 N개의 샘플이고 입력 음성 스트림이 x_0, x_1, \dots 이라면 패킷 P_1 은 짝수 샘플 $x_0, x_2, \dots, x_{2N-2}$ 이고 패킷 P_2 는 홀수 샘플 $x_1, x_3, \dots, x_{2N-1}$ 이라고 가정하자. 만약 수신 측이 P_2 를 제외한 P_1 을 수신한

다면 그때 P_2 는 다음과 같이 복원된다.

$$\hat{x}_{2j-1} = \frac{x_{2j-2} + x_{2j}}{2} \quad \text{where } j=1, \dots, N \quad (2.1)$$

1. Adaptive-Filter를 적용한 복원 방법

적용 시스템은 특정 입력에 부응하여 특별한 참조 출력을 발생시키기 위한 "reference signal" 효과에 적합한 continuous-range 변수들을 갖는 장치이다[8]. 적응 과정은 실제로 다음 특성을 갖는 최적의 과정이다 : (a) 최적 변수들의 이전을 알지 못한다고 가정하자. (b) 적응 필터는 최적 변수에 도달하기 위해 스스로 변함없이 적응할 수 있다. 그림 1에서 일반적인 Adaptive Filter를 보인다.

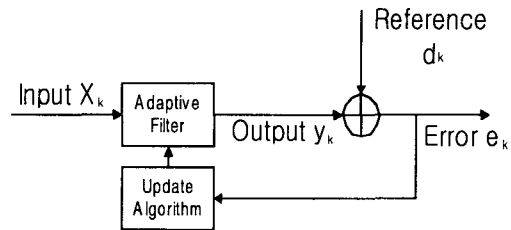


그림 1. An Adaptive filter

일반적으로 입력 신호로부터 압축된 정보를 사용하는 비적용 필터들은 통상적으로 선형 시간에 불변이다. 즉, 시간에 개되지 않고 모든 신호에 대한 선형 신호의 동일한 설정을 정확하게 실행한다[9]. 적응 필터의 경우 시간 변화에 대한 제한이 없어진다. 이것은 임의로 미리 설정된 최적 표준에 의해 선형 필터링 작용에 사용되는 상관 계수를 바꾸는 필터를 허용함으로써 실행된다. 만약 우리가 이전 최적 필터 상관계수를 알지 못한다면 적응 필터는 단계적으로 최적의 상관계수를 찾으려는데 사용될 수 있다.

간단한 LMS(Least Mean Square) 적응 필터 알고리즘은 다음처럼 요약 될 수 있다.

$$y_k = \sum_{l=0}^{M-1} w * x_{k-l} \quad (2.2a)$$

$$e_k = d_k - y_k \quad (2.2b)$$

$$\hat{\nabla}^k = \frac{\partial e_k^2}{\partial \vec{w}} = -2e_k \vec{x}^k \quad (2.2c)$$

$$\vec{w}^{k+1} = \vec{w}^k - \mu \hat{\nabla}^k = \vec{w}^k + 2\mu e_k \vec{x}^k \quad (2.2d)$$

초기 계수 $\vec{w} = 0, \dots, 0, 0.5, 0, 0.5, 0, \dots, 0$ 이고 평균 복원 신호를 이용하여 참조 신호로 선택한다.

III. Transformation-Based Reconstruction

이 장에서는 II장에서 소개한 Adaptive-Filter-Based 방법을 interleaving과 평균 복원을 기본으로 하여 송신 측에 적용하는 transformation-Based 복원 알고리즘을 소개한다. 송신 측의 원 데이터 스트림 $\vec{x} = x_0, x_1, \dots, x_{2N-1}$ 는 식(3.1)에 의해 $\vec{y} = y_0, y_1, \dots, y_{2N-1}$ 스트림으로 된다. averaging 후의 스트림에 (2.2a) - (2.2b)식을 적용하여 변형 스트림 $\vec{y} = \hat{y}_0, \hat{y}_1, \dots, \hat{y}_{2N-1}$ 이 산출된다.

$$y_i = \begin{cases} \frac{x_{i-1} + x_{i+1}}{2}, & i \text{ even and } i \neq 0 \\ \frac{x_{i-1} + x_{i+1}}{2}, & i \text{ odd and } i \neq 2N-1 \\ \frac{x_{i+1}}{2}, & i = 0 \\ \frac{x_{2N-2}}{2}, & i = 2N-1 \end{cases} \quad (3.1)$$

수신 측에서는 변형된 신호를 수신한 후 averaging을 사용하여 손실된 패킷을 복원한다.

변형 방법을 사용하여 손실이 발생했을 경우 더 좋은 복원율을 허용하기 위해 우리는 반복적인 two-way interleaving에 의해 M-way interleaving을 만든다. 예를 들어 다음과 같은 four-way interleaving을 생각할 수 있다[10].

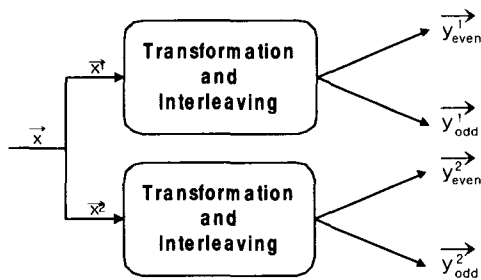


그림 2. two-way interleaving을 사용하여 four-way interleaving 구성

그림 2는 송신 측에서 만들어진 4개의 스트림을 보인다. 송신 측에서 먼저 원 스트림 \vec{x} 을 $\vec{x}^1 (= x_0, x_2, \dots)$ 과 $\vec{x}^2 (= x_1, x_3, \dots)$ 으로 interleave한다. 다음에는 2장에서 논한 변형과정

을 사용하여 \vec{x}^1 과 \vec{x}^2 를 변형하고 \vec{y}^1_{even} , \vec{y}^1_{odd} , \vec{y}^2_{even} 그리고 \vec{y}^2_{odd} 의 4개의 스트림을 얻기 위해 2개의 변형 스트림을 각각 interleave 한다.

수신 측에서 손실이 없을 때 \vec{y}^1_{even} 과 \vec{y}^1_{odd} 을 deinterleaving 하여 de-interleaved 스트림에 inverse 변형을 제공함으로써 \vec{x}^1 을 구성할 수 있다. 유사하게 \vec{x}^2 는 \vec{y}^2_{even} 과 \vec{y}^2_{odd} 로부터 복원될 수 있다.

임의의 패킷들이 수신 측에서 손실되었을 때, 4가지 가능성이 있다.

첫 번째의 경우 interleaving 세트에서 3개 패킷들은 손실되어진다. 예로 \vec{y}^1_{odd} , \vec{y}^2_{even} 그리고 \vec{y}^2_{odd} 이 손실되었다고 가정하자. Averaging 방식에 의해 수신된 \vec{y}^1_{even} 으로부터 \vec{y}^1_{odd} 를 복원할 수 있다. \vec{x}^1 을 복원한 후 인접 짝수 샘플들의 평균을 계산함으로써 \vec{x}^2 의 각 샘플을 복원할 수 있다. 이러한 경우 \vec{y}^1_{odd} 가 최적으로 복원되기 때문에 복원 질은 변형이 없을 때보다 더 좋게 보장되어진다.

두 번째로 2개의 패킷들에서 \vec{x}^1 나 \vec{x}^2 패킷이 분실되었을 경우이다. \vec{y}^2_{even} 과 \vec{y}^2_{odd} 이 분실되었다고 가정하자. 우리는 \vec{x}^1 을 얻기 위해 \vec{y}^1_{even} 과 \vec{y}^1_{odd} 에 제안된 변형을 적용할 수 있다. 그리고 평균에 의해 \vec{x}^2 를 복구한다.

세 번째로 \vec{x}^1 의 한 패킷과 \vec{x}^2 의 한 패킷, 두 패킷이 손실되었을 경우이다. 제안된 방식으로 2개의 손실된 패킷들을 수신된 패킷들을 기준으로 최적으로 복원된다. 이 경우도 변형이 없는 경우보다 더 좋은 실행을 보장할 수 있다.

마지막의 경우 interleaving 세트의 단 한 개의 패킷이 손실되었다. 예로 \vec{y}^2_{odd} 가 손실되었다고 가정하자. 우리의 방법을 사용하여 우리는 수신된 \vec{y}^2_{even} 로부터 최적의 \vec{y}^2_{odd} 을 복원할 수 있다. 우리는 \vec{y}^1_{even} 과 \vec{y}^1_{odd} 에 우리의 inverse 변형을 제공함으로써 \vec{x}^1 을 복원할 수 있다.

IV. 실험 결과

두 개의 복원 방법을 기본으로 홀수 샘플들이 수신 측에서 손실되었고 짝수 샘플들이 잃어버린 샘플들을 복원하는 사용되었다고 가정하자. 그림 3은 잃어버린 홀수 샘플이 두 개의 인접 짝수 샘플

플들의 평균으로 계산되어지는 것과 두가지 복원 방법과의 결과를 보여준다. averaging에 적응 필터링을 적용한 방법에서 잃어버린 샘플들은 averaging을 사용하여 먼저 복원되고 그때 스트림은 적응 필터를 통과한다. 적응 필터링 후 복원된 스트림은 적응 필터링이 파형의 모양을 추적할 수 있기 때문에 원 스트림의 근사치보다 더 좋다. 그러나 적응 필터링은 그림에서 보이는 것처럼 원 신호의 빠른 변화에는 적용할 수 없다. transformation-based 방법에 의해 복원된 데이터는 송신 측에서 변형을 하고 변형된 짝수 샘플로 홀수 샘플이 복원이 된다. 그림에서 변형 샘플들을 기본으로 한 복원 스트림이 원 스트림에 더 근접한 것을 볼 수 있다.

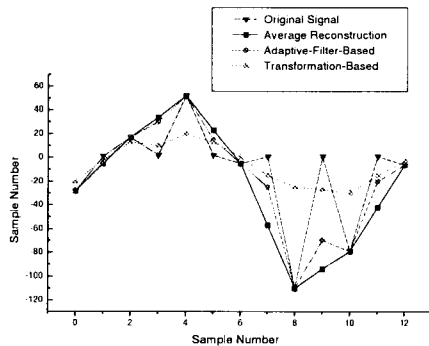


그림 2. Adaptive-Filter Based와 Transformation-Based 비교 결과

V. 결 론

이 논문에서 인터넷상의 실시간 음성 전송을 위해 두 가지 방법을 저비트율 코덱에 적용해 보았다. 첫 번째 적응 필터링을 적용한 복원에서는 빠르게 바뀌는 음성 신호를 다루는데 어려움을 겪었다. 송신 측에서 신호를 변형하는 방법은 빠르게 바뀌는 음성 신호에서 잘 실행된다. 우리는 앞으로 송신 측에서 신호를 변형할 때 더 정확한 전송 에러를 예측하여 원 신호에 더 근접하게 복원할 수 있도록 초점을 맞춘다.

참고문헌

[1] 지경용, 고중걸, "각 국가별 차세대 인터넷 동향분석, 주간기술동향 통권 987호, pp.22, march, 2001.
 [2] J. Suzuki and M. Taka, "Missing packet recovery techniques for low-bitrate coded speech", IEEE J.select. Areas Commun., vol. 7, 707-717, June 1989.

[3] O. J. Wasem, D. J. Goodman, C. A. Dvordak, and H. G. Page, "The effect of waveform substitution on the quality of PCM packet communications", IEEE Trans. Acoust., Speech, Signal Processing, vol. 36, pp. 342-348, Mar. 1988.
 [4] V. Hardman, M. A. Sasse, M. Handley, and A. Watson, "Reliable audio for use over the internet", in Proc. Int. Networking Conf., June 1996, pp. 171-178.
 [5] Telogy Networks, <http://www.webproforum.com/> Int, Eng. Consortium Web Proforums, Voice over packet tutorial, Nov. 1997.
 [6] N.Shacham and P.Mckenney, "Packet recovery in high-speed networks using coding and buffer management", int Proc. IEEE INFOCOM, May 1990, pp. 124-131.
 [7] J. C. Bolot and P. Hoschka, "Adaptive error control for packet video in the Internet", in Proc. Int, Conf, Image Processing, Sept, 1996, vol. 1, pp. 25-28.
 [8] O. Macchi, Adaptive Processing: The Least Mean Squares Approach with Applications in Transmission. Chichester, England: John Wiley & Sons, 1995.
 [9] C. Cowan and P.Grant, Adaptive Filters, Englewood Cliffs, NJ: Prentice-Hall, 1985.
 [10] Benjamin W. Wah and Dong Lin, "Transformation-Based Reconstruction for Real-Time Voice Transmissions over the Internet", IEEE TRANSACTIONS ON MULTIMEDIA, VOL.1, NO. 4, DECEMBER 1999