

상태별 이득 제어 및 fixed codebook estimation을 이용한 G.729에서의 Packet Loss Concealment 알고리즘 개선

문 광, 한 민 수
한국정보통신대학원대학교 음성/음향정보연구실

Improvement of Packet Loss Concealment Algorithm by Using state gain control and fixed codebook estimation

Kwang Moon, Minsoo Hahn
Speech and Audio Information Lab., Information and Communications University
E-mail: mlight@icu.ac.kr

Abstract

In real time packetized voice applications, missing frames is a major source of voice quality degradation. Thus packet loss concealment(PLC) algorithms are needed to guarantee the QoS of the VoIP. Still current speech codecs for VoIP work poor when consecutive packet losses are issued. In this paper, we proposed a new PLC algorithm for the G.729 codec. Our algorithm works better especially when the consecutive packet loss occurs mainly because it adopts an adaptive gain controller utilizing the number of missing packet information combined with a fixed codebook vector estimation algorithm and LPC bandwidth expansion.

I. 머리말

차세대 통합망의 도래와 함께 지금까지의 회선 기반의 통화망이 아닌 패킷 기반의 통화망으로의 진화가 이루어지면서 주목 되는 것이 패킷 기반의 전화망인 VoIP(Voice Over IP)이다. 그러나 IP 네트워크는 트래픽이 증가함에 따르는 끊어짐(connectionless) 현상으로 인하여 낮은 통화품질을 나타내며 VoIP의 통화품질을 저하시키는 요인 중 가장 큰 것이 패킷 손실이다. 패킷 손실은 네트워크의 트래픽이 증가함에 따라 실시간 전송을 요구하는 VoIP에서 패킷이 도착하지 못 하여 패킷이 재생되지 못 하는 것을 말하며, 이러한 패킷 손실은 실제 통화 품질에 매우 큰 영향을 미치게 된다[3].

현재 국내 VoIP 통화망에서 사용되는 코덱은 G.711, G.723.1, G.729 등의 세 가지 코덱이 대표적이며 그 중에서 낮은 비트 전송율을 유지하면서도 32 bit ADPCM (G.711) 수준의 고품질의 음성 품질(toll quality) 유지하는 코덱이 G.729 이다. G.729는 비트율(bit rate)이 작기 때문에 한정된 자원을 효율적으로 사용할 수 있으며[5], 높은 수준의 품질을 보장하기 때문에 VoIP 코덱들 중에서 중요한 위치를 차지하고 있다. G.729를 비롯한 모든 코덱들은 자체적으로 패킷 손실 은닉 알고리즘을 내장하고 있다. 본 논문은 이러한 효율적인 품질을 나타내는 G.729 코덱의 연속적인 패킷 손실에 대한 패킷 손실 은닉알고리즘을 개선함으로써 보다 향상된 통화품질을 제공하려 한다.

II. 본론

1. G.729에서의 프레임 손실 은닉 알고리즘

프레임 손실이 발생하면 해당 프레임의 LP 계수, 피치 정보, 고정 코드북 이득 및 적응형 코드북 이득 등의 정보가 손실된다. G.729 음성 코덱에서는, 일단 프레임 손실이 검출되면, 바로 직전의 양호한 프레임의 스펙트럼 파라미터들을 이용하여 새로운 파라미터들을 생성한다[5]. LP 계수는 이전 프레임에서 사용한 LP 계수를 손실된 프레임에 대해 반복적으로 사용하고, 피치의 경우는 이전 서브 프레임의 피치 값에 1을 더하여 사용한다. 적응형 코드북 이득(adaptive codebook gain) 및 고정 코드북 이득(fixed codebook gain) 값은 이전 프레임의 값에 감쇠 인자를 곱하여 손실된 프레임에서 사용한다. 손실된 프레임의 음성코딩 파라미터들이 기본적으로 이전 값과 약간 다르거나 혹은 이전 값의 감쇠된 형태로 사용하는 이유는 울려 퍼지는 음

(reverberant sound)을 방지하기 위해서이다.[2]

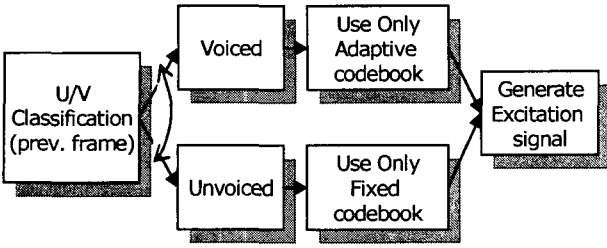


그림 1. G.729 PLC 알고리즘의 블록 다이어그램

2. 제안한 알고리즘

2.1 개요

G.729에서의 패킷 손실 은닉 알고리즘의 특징은 그림 1에서와 보는 바와 같이 손실되기 이전 프레임이 유성음인지 무성음인지를 판단하여 손실된 프레임에 대해 각기 다른 코드북 벡터를 사용한다. 이전 프레임이 유성음으로 판단된 경우에는 적응형 코드북 벡터만을 사용하고 무성음으로 판단된 경우에는 고정 코드북 벡터만을 사용한다[5]. 또한 CELP 계열의 코덱에서는 하나의 프레임이 손실되면 여기신호(excitation signal)가 손실된 프레임 이후의 갱신된 메모리의 내용에 영향을 미친다[2]. 본 논문에서 제안하는 알고리즘의 주요 개념은 그림 2.에서 보듯이 이득값 조절과 고정 코드북 벡터를 추정하는 절차에 관계된 것이다. 본 논문에서는 두 가지 코드북 벡터 모두를 사용하며 손실된 패킷 개수에 따른 이득제어를 사용하여 적응형 코드북 이득 값을 조절한다. 아울러 두 가지 코드북 벡터를 모두 사용하기 위해서는 고정 코드북 벡터에 대한 추정 알고리즘이 요구된다.

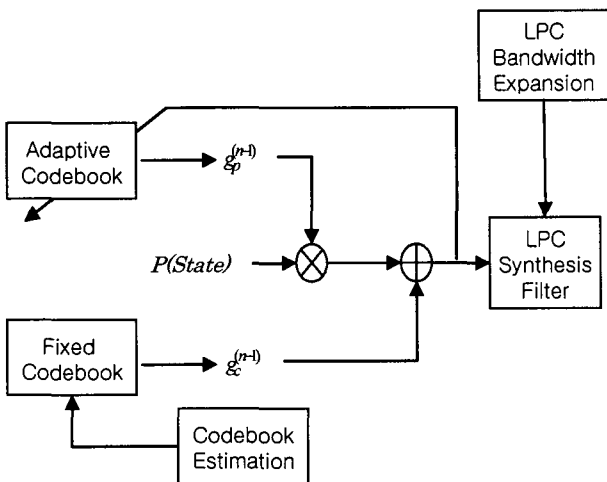


그림 2. 제안한 알고리즘의 블록 다이어그램

그 외에도 대역폭 평활화를 이용하여 음질을 향상시킨다.

2.2 상태별 이득제어

G.729에서는 프레임이 손실된 경우 적응형 코드북 이득 및 고정 코드북 이득 값을 식 (1)에서와 같이 이전 프레임의 이득 값에 미리 정해진 감쇠 인자 0.98을 곱하여 감쇠된 값을 사용한다[5].

$$g_n = 0.98g_{n-1} \quad (1)$$

그러나 이러한 방식은 프레임이 손실된 이후 양호한 프레임(good frame)이 도착하여도 양호한 프레임으로 에러가 전파되는 에러 전파(error propagation) 현상이 나타나게 된다. 프레임이 무작위로 1개씩 손실된 경우에는 에러 전파 현상이 큰 영향을 미치지 않지만 연속적인 손실의 경우에는 에러 전파 현상이 음질에 큰 영향을 미친다. 이러한 에러 전파 현상을 줄이기 위해 적응형 코드북 이득값을 조절해 주어야 한다[4].

$$g_p^n = P(state) \cdot g_p^{n-1} \quad (2)$$

본 논문에서는 이러한 이득 값 조절을 위하여 식 (2)에서와 같이 상태별 이득 제어를 수행한다. 여기서 상태(state)란 연속적으로 손실된 프레임의 개수를 말하며 연속적인 프레임 손실의 경우 이러한 상태에 따라 P(state) 인자를 이용하여 이득 값을 조절해 줌으로써 에러 전파 현상을 줄일 수 있다. 표 1.은 각 상태에 따른 이득값(P(state))을 나타낸다. 표 1. 과 같이 이득 값을 정한 이유는 변화하는 음성의 특성을 반영하고 에러 전파를 방지하기 위해서이다.

2.3 fixed codebook vector 추정

G.729에서는 그림 1.에서 보는 바와 같이 이전 프레임이 유성음인지 무성음인지에 따라 사용하는 코드북 벡터를 달리한다. 이전 프레임이 무성음으로 판명된 경우 사용하는 고정 코드북 벡터는 랜덤하게 4개의 펄스 위치 정보를 얻은 후 이전 프레임의 고정 코드북 이득 값과 피치 정보를 이용하여 코드북 벡터를 생성한다[5].

표 1. 상태별 gain control 인자

State	P(state)
1	1.2
2	1.1
3	1.0
4	0.9
5	0.8
6	0.7

이러한 방법은 무성음 구간에서는 매우 잘 동작하게 된다. 그러나 유성음 구간에서도 고정 코드북 여기신

호(fixed codebook excitation signal)에는 주기적 구조(periodicity structure)가 여전히 있다. 고정 코드북 벡터를 단순히 0으로 채워넣는 것은 이러한 주기적 구조를 사용하지 못하게 되는 것이다. 게다가 이러한 구조가 서로 연관이 있는 ‘펄스 위치’ 및 ‘이득’의 형태로 하나의 패턴을 이룬다[2]. 본 논문에서는 적응형 코드북 벡터와 고정 코드북 벡터를 모두 사용하며, 손실된 프레임과 인접한 프레임들을 조사하여 손실된 고정 코드북 벡터를 추정하여 복원한다. 본 논문에서 손실된 프레임의 고정 코드북 벡터를 추정하기 위해 사용한 방법은 직전 프레임에서의 고정 코드북 이득 값을 이용하여 이전 6개의 서브프레임에서 가장 큰 값을 갖는 고정 코드북 이득 값을 찾은 후 찾아낸 서브 프레임의 이후 서브프레임의 고정 코드북 벡터를 손실된 프레임의 고정 코드북 벡터로 사용하였다.

2.4 bandwidth expansion

G.729에서는 LP 계수의 경우 이전 프레임에서의 LP 계수를 손실된 프레임에 반복적으로 사용한다.[4] 본 논문에서는 손실된 구간의 LP계수를 이전 프레임의 LP계수를 반복적으로 사용하지 않고 식 (2)를 이용하여 대역폭 확장을 통한 LP계수를 사용한다[5].

$$a_i' = (0.98)^i a_i \quad i=1, \dots, p \quad (3)$$

식 (3)에서 a_i 와 a_i' 는 이전 프레임의 LP 계수 및 현재 프레임에서의 LP 계수를 나타낸다. LP필터의 대역폭을 확장한다는 것은 필터의 피크를 약간씩 줄이고 대역폭을 조금 늘리는 것을 의미한다. 이전 프레임에 가파른(sharp) 포먼트가 있는 경우 이전 프레임의 LP 계수를 반복해서 사용하면 인공음(artifact) 같은 음질이 나와 통화 품질을 저해하게 된다[4]. 이런 경우 이전 프레임으로부터 얻은 LP 계수를 가지고 대역폭 평활화(bandwidth expansion)를 통하여 조금 더 부드러운 형태의 LP필터를 생성한다.

이로 인해 천천히 변화하는 음성신호의 스펙트럼 특성을 반영할 수 있으며 이러한 LP계수를 이용한 음성합성으로 보다 부드러운 음성 변화구간을 얻을 수 있다.

III. 실험 및 결과

실험은 10 개의 문장 셋(남 4 여 6)에 대하여 각각의 손실을 별로 연속적으로 프레임 손실을 유발 시킨 후 디코딩된 신호에 대해 9명의 청취자에게 선호도 조사를 수행하였다. 두 실험 문장 사이에 매우 큰 차이가 나지 않으므로 점수를 기록하는 MOS 방법보다는 주관적 평가 방법이지만 두 데이터 사이의 우위를 가릴 수 있는 A/B 선호도 조사(A/B preference test)를

수행하였다[4].

표 2. 프레임 손실율에 따른 선호도 (2연속 손실)

	Strong	Slight	Equal	Slight	Strong
	기존 알고리즘			제안 알고리즘	
3%	0	15	56	19	0
5%	0	11	33	43	3
10%	0	8	40	33	9

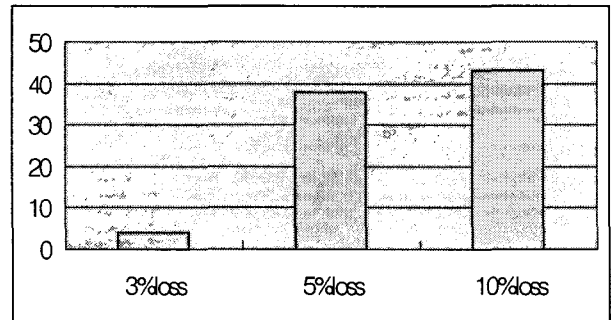


그림 3 가중치를 이용한 표 2의 그래프 표현

표 3. 프레임 손실율에 따른 선호도 (3연속 손실)

	Strong	Slight	Equal	Slight	Strong
	기존 알고리즘			제안 알고리즘	
3%	0	14	58	18	0
5%	0	19	36	33	2
10%	0	16	44	25	5

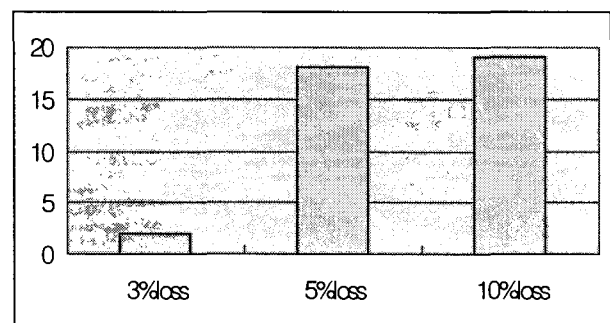


그림 4 가중치를 이용한 표 3의 그래프 표현

실험과정은 실험 대상이 되는 10 개의 실험용 음성을 G.729로 인코딩을 한 후 인코딩된 상태에서 패킷 손실율에 따라 연속적으로 패킷 손실을 유발 시킨 다음 디코딩 한 신호를 실험에 사용하였다. 손실율은 각각 3% 5% 10% 등으로 정하였다. 손실율이 10%가 넘어가면

청취자가 받아들일 수 있는 음성 품질을 넘어가게 되기 때문에 10%를 최대 손실율로 정하였다. 실험에 대한 수행 결과는 표 2와 표 3에서 보는 바와 같다. 그림 3과 그림 4는 표 2와 표 3에서의 결과를 가중치를 적용해서 나타낸 그래프적 표현이다. 연속적으로 손실된 패킷의 수는 2개 및 3개의 연속 손실을 랜덤하게 발생시켰다. 그림 3과 그림 4는 표 2와 표 3의 그래프 표현이다. 그림 3과 표 2에서 보듯이 연속적으로 손실된 프레임의 수가 2인 경우 손실율이 3%인 경우보다 5% 이상에서 실험에 사용한 알고리즘이 기존의 알고리즘보다 더 나은 성능을 보인다. 그림 4와 표 3의 3연속 패킷 손실의 경우는 2연속 패킷 손실의 경우보다는 그 차이가 약간 줄어들긴 하였으나, 기존의 G.729에 내재된 패킷 손실 알고리즘에 비하여 더 나은 성능을 보임을 알 수 있다. 표 2와 표 3을 비교해 보았을 때, 2연속 손실이었을 때보다는 성능 개선의 폭이 조금 작지만 기존의 G.729 표준 알고리즘 보다는 나은 성능을 보임을 알 수 있으며, 따라서 본 논문에서 제안하는 알고리즘의 성능이 연속적 패킷 손실에서 G.729 코덱에서 제공하는 알고리즘보다 기능이 개선되었음을 알 수 있다.

IV. 결론

본 논문에서는 G. 729에 내장되어 있는 패킷 손실 알고리즘에서 연속적 패킷 손실에 대해 효과적인 패킷 손실 알고리즘을 제안하였다. 제안된 알고리즘의 특징 중 첫 번째는 연속적으로 손실된 프레임의 수에 따른 이득값 조절과 두 가지 코드북 벡터의 동시 사용과 고정 코드북 벡터의 추정 알고리즘, 그리고 부드러운 음성구간 합성을 위한 대역폭 평활화 등이다. 연속적이 아닌 단일 프레임 손실의 경우에는 G.729 표준안에서 사용하는 패킷 손실 은닉 알고리즘을 사용해도 커다란 음질 저하를 느낄 수 없으나 프레임이 연속적으로 손실된 경우에는 이전 프레임에서의 파라미터를 반복하기 보다는 손실된 패킷의 수에 따라 이득값을 조절해주는 적응형 이득 제어(adaptive gain controller) 방식을 통하여 패킷 손실에 따른 음질 저하를 효과적으로 방지 할 수 있었다. 고정 코드북 벡터 추정 및 대역폭 평활화 방법은 앞서 설명한 이득값 제어 알고리즘과 함께 사용함으로써 음성 품질의 향상을 꾀할 수 있다. 손실율이 작은 경우에는 손실 구간이 짧아서 두 데이터 사이의 차이가 크지 않으나 손실율이 커짐에 따라 손실구간이 늘어나면 기존의 알고리즘 보다 제안된 알고리즘의 성능이 나아짐을 알 수 있다.

참고문헌

- [1] Hong Kook Kim, Hong-Goo Kang, "A frame erasure concealment algorithm based on gain parameter re-estimation for CELP coders" *Signal Processing Letters, IEEE* , Volume: Sept. 2001 Page(s): 252-256
- [2] Jian Wang and Jerry D.Gibson, "Parameter interpolation to enhance the frame erasure robustness of CELP coders in packet networks" *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 001 IEEE International Conference on* , Volume: 2, 7-11 May 2001 Page(s): 745-748 vol.2
- [3] Anton Kos, Borut Klepec, Saso Tomazic, "Techniques for performance improvement of VoIP applications" *Electrotechnical Conference, 2002. MELECON 2002. 11th Mediterranean*, 7-9 May 2002 Page(s): 250-254
- [4] Juan C.Martin, Takahiro Unno, Vishu Viswanathan, "Improved frame erasure concealment for CELP-based coders", *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on* , Volume: 3, 5-9 June 2000 Page(s): 1483-1486 vol.3
- [5] Salami, R.; Laflamme, C.; Adoul, J.-P.; Kataoka, A.; Hayashi, S.; Moriya, T.; Lamblin, C.; Massaloux, D.; Proust, S.; Kroon, P.; Shoham, Y.; *Design and description of CS-ACELP: a toll quality 8 kb/s speech coder* *Speech and Audio Processing, IEEE Transactions on* , Volume: 6 March 1998 Page(s): 116-130