

論文2002-39SP-6-13

# VoIP 손실 환경에 강인한 저지연 LSF FEC 기법 (Low-Delay LSF FEC Technique Robust in Lossy VoIP Environment)

梁海龍\*, 李垌勳\*, 黄仁浩\*

(Hae-Yong Yang, Kyung-Hoon Lee, and In-Ho Hwang)

## 요약

VoIP 음성 패킷 손실에 대한 대응 방안으로 제시되고 있는 매체 종속 FEC 기법은 통화 품질을 개선시키는 효과를 갖는데 반하여 한 프레임에 해당하는 추가지연이 발생하는 단점을 갖는다. 본 논문에서는 패킷 손실 복원에 사용되는 잉여 정보로 미래 프레임의 LSF 성분을 사용함으로써, 전송 지연을 줄이고 통화 품질을 개선할 수 있는 LSF FEC 기법을 제안하고 그 성능을 평가한다. 성능 평가를 위해서 VoIP에서 사용하는 ITU-T G.723.1, G.729 코덱을 Gilbert 손실 모델에 적용하고, PESQ 음질 측정 알고리즘을 사용하여 각 손실을 별로 MOS를 추정하는 방법을 사용한다. 본 논문에서 제안한 기법은 기존의 매체 종속 FEC 기법에 비해서 6.5ms~27ms 이상의 지연 감소 효과를 가지고 있는 것으로 나타났으며, FEC를 적용하지 않은 경우와의 복원 음성 품질 비교 시뮬레이션 결과, 5% 정도의 현실적인 손실 환경에서 MOS 0.1 이상의 음질 개선 효과를 보였다.

## Abstract

Media-specific FEC techniques, suggested to confront with VoIP speech packet loss, improve speech quality at the expense of generating additional one-frame delay. In this paper, we suggest new media-specific FEC, i.e. LSF FEC technique which is able to improve speech quality with much shortened additional delay. In the proposed technique, the LSF parameters of the future frame are utilized to recover a lost packet. To evaluate performance of the proposed technique, we use ITU-T G.723.1 and G.729 Codec and apply Gilbert packet loss model and estimate MOS per every packet loss rate using PESQ speech quality estimation algorithm. The proposed technique has effect of shortening delay over from 6.5ms to 27ms compared with existing media-specific FEC techniques. Simulation results for comparison of reconstructed speech quality show this novel technique improves the MOS over 0.1 in practical lossy environment of 5% packet loss rate.

**Keywords** : VoIP(Voice over IP), FEC(Forward Error Correction), LSF(Line Spectral Frequency)

## I. 서론

인터넷 트래픽의 급격한 증가와 서비스 보급의 대중

화에 힘입어 통화 요금 절감 및 네트워크 통합의 차원에서 인터넷 전화가 전 세계 통신 및 인터넷 서비스 사업자들의 주요 관심사가 되고 있다.

인터넷 전화의 근간이 되는 VoIP 기술은 실시간 음성 신호 전송을 지원하기 위해서, 데이터 손실 시 복구 메커니즘을 내장한 TCP 대신 손실 가능성을 내포한 UDP 프로토콜을 사용한다. 이는 손실보다 지연이 통화 품질에 더 치명적인 영향을 주는 음성 신호의 특징으

\* 正會員, 國家保安技術研究所 應用技術研究部

(Application Technology Research Department, National Security Research Institute)

接受日字:2002年7月15日, 수정완료일:2002年10月15日

로 인해 손실을 감수하고 저지연 실시간 전송을 실현하기 위함이다. 따라서 VoIP에서 음성 데이터의 손실은 데이터 네트워크에서의 음성 전송에 수반되는 피치 못할 현상이라 할 수 있지만, 손실 역시 통화 품질에 많은 영향을 미치는 것으로 나타났으며,<sup>[11]</sup> 이 점이 현재 VoIP 대중화의 가장 큰 걸림돌로 작용하고 있다. 이에 따라 실시간 전달 특성을 갖는 멀티미디어 서비스 시스템에서의 전송 손실에 따른 QoS 보장 방안이 VoIP 관련 분야에서 큰 이슈가 되고 있으며, 이를 위한 많은 연구가 진행되고 있다.

본 논문에서는 현재 제시되고 있는 여러 가지 QoS 보장 방안들 중에서 매체 종속 FEC(media-specific forward error correction) 기법<sup>[2]</sup>을 개선하여, 잉여 정보 전송에 미래 프레임의 LSF 성분을 사용함으로써 전송 지연을 줄이고 통화 품질을 개선할 수 있는 저지연 LSF FEC 기법을 제시하고 그 성능을 평가한다. 이어지는 2장에서는 VoIP에서 사용되는 기존의 음성 프레임 손실 복구 기법을 소개하고, 3장에서는 본 논문에서 제안하는 저지연 LSF FEC 기법에 대해서 설명한다. 4장과 5장에서는 제시한 기법에 대한 성능 평가 방법과 시뮬레이션 결과에 대해서 논하고 6장에서 결론을 맺는다.

## II. 음성 프레임 손실 복구 기법

현재 대표적인 VoIP QoS 보장 대책으로 제시되고 있는 것으로는 DiffServ, RSVP와 같은 네트워크 프로토콜을 이용한 네트워크 측면의 접근법과, 단말 측에서 음성 패킷을 조작하는 단말 측면에서의 접근법이 있다. 네트워크 측면에서의 접근법은 QoS 보장을 위한 궁극적인 대책이기는 하지만 전체 네트워크에 적용하기에는 시기상조이기 때문에 지금까지는 단말 측면에서의 QoS 보장 기법이 주로 사용된다.

단말 측 QoS 관리 대책은 크게 송신 단말 기반(sender-based)과 수신 단말 기반(receiver-based) 방식으로 구분할 수 있다. 송신 단말 기반 방식은 송신 단말이 QoS 보장을 위해서 송신 패킷에 적절한 처리를 수행하고 수신 단말이 복원하는 방식을 말하며, 수신 단말 기반 방식은 수신 단말이 손실된 패킷을 음성 신호의 특성을 이용해서 재생해 내는 방법으로, 일명 오류 은닉(error concealment) 기법이라고도 부른다. 송신 단말 기반 방식은 대역폭, 지연 등이 증가하는 단점이

있는 반면, 수신 단말 기반 방식은 손실률이 커질 경우에는 복원 성능이 떨어지는 단점이 있다. 현재는 이 두 기법을 동시에 사용하는 것이 성능이 가장 좋은 것으로 알려져 있다.<sup>[2]</sup>

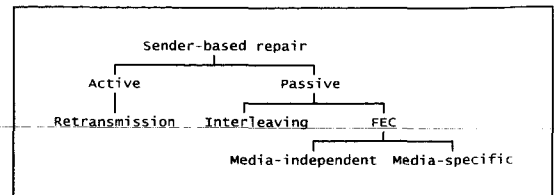


그림 1. 송신 단말 기반 복원 방식 분류

Fig. 1. Taxonomy of sender-based loss repair method.

송신 단말 기반 방식은 <그림 1>과 같이 분류가 가능하며<sup>[2]</sup> 이 중에서 FEC 기법은 전송 데이터 이외의 잉여 정보를 추가 전송함으로써 손실이 발생했을 경우 추가한 잉여 정보를 이용해서 복원하는 기법이다. 이 기법은 매체의 성질과 무관한 매체 독립(media-independent) 기법과 매체 종속(media-specific) 기법으로 분류가 가능하다. 매체 종속 FEC는 오류 복구에 사용되는 잉여 정보를 전송할 매체(음성, 영상, 데이터)의 종류에 따라 설정하는 방법으로, 매체의 특성과 관련이 있기 때문에 매체 독립 기법에 비해 적은 잉여 정보를 사용하면서도 좋은 복원 음성 품질을 나타낸다. VoIP에서 전송하는 음성 매체에 대해서 제시되고 있는 잉여 정보 추출 알고리즘으로는 short-term 에너지와 영교차율(zero crossing ratio)을 사용한 방법<sup>[3]</sup>과 이중 코덱(dual codec)을 이용한 방법이 있다.<sup>[4,5]</sup> [4,5]에서는 주 코덱(primary codec)으로는 full rate GSM 코덱(13.2kbps)을, 잉여 정보를 만드는 부코덱(secondary codec)으로는 LPC 코덱(2.4~5.6kbps)을 사용해서 매체 종속 FEC를 구현한다. <그림 2>는 이중 코덱 방식의 손실 복구 기법을 설명하기 위한 부호화, 복호화 흐름을 나타낸다. 이 방식은 GSM 고 전송률 코덱, LPC 저 전송률 코덱을 이용해서 각각 부호화를 수행하고, 전송 시에는 n번째 프레임의 고 전송률 음성 패킷과 n-1 프레임의 저 전송률 음성 패킷을 잉여 정보로 묶어서 전송한다. 만일 n+1 패킷에 손실이 발생하였을 경우에 수신 측에서는 n번째 패킷에 포함되어 있는 저 전송률 음성 패킷을 이용해서 복원을 수행한다.

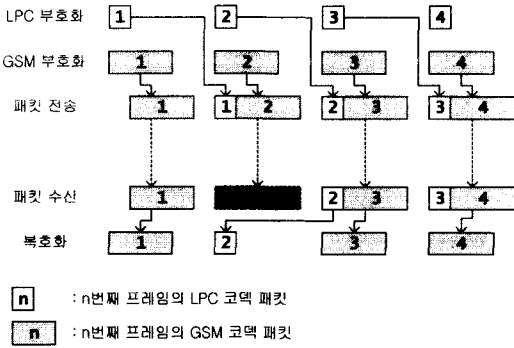


그림 2. 이중 코덱 방식의 손실 복구 흐름  
 Fig. 2. Loss repair flow of dual codec method.

이와 같은 매체 종속 FEC는 손실 환경에서 복원률이 높아서 좋은 통화 품질을 제공하지만 한 프레임의 추가 지연이 발생하는 단점이 있다. 매체 종속 FEC 기법의 지연 분석을 <그림 3>의 타이밍도를 이용해서 설명한다.

① 1 번째 입력 음성 프레임은 프레임 지연(DF)과 lookahead 지연(DL) 후에 압축을 시작하고, 압축 처리 지연(DEP) 후에 음성 패킷 P(1)을 생성한다. 이 때 코덱에 의한 알고리즘 지연(DA)은 프레임 지연과 lookahead 지연의 합으로 나타난다. ( $D_A = D_F + D_L$ )

② 1 번째 음성 프레임 손실을 대비한 잉여 정보 R(1)을 추출한다. 잉여 정보는 코덱 압축에 비해서 상대적으로 적은 복잡도 및 데이터 량을 가지는 것이 일반적이다. 이중 코덱 방식의 경우 잉여 정보는 LPC 코

덱 음성 패킷이 된다.

③ 압축 지연( $D_{EMFEC}$ ) 후에 1 번째 패킷을 전송한다. 압축 처리 지연이 적으면 전체 지연이 감소하지만 이를 위해서는 하드웨어 및 소프트웨어 비용이 지연 감소에 비해 급격히 증가하고 입력 음성 신호의 종류에 따른 수신 패킷 지터(jitter) 발생 원인이 될 수 있다. 따라서 한 프레임 시간 내에 압축을 수행함으로써 실시간 압축 특성을 보장하면서 다음 프레임의 시작 부분에서 전송하는 방법이 현실적으로 많이 사용된다. 1 번째 프레임에 대한 전송 데이터는  $[R(0)|P(1)]$ 으로 나타난다. 이 때 압축 지연은 식 1과 같이 나타난다.

$$D_{EMFEC} = D_A + D_{EP} \quad (1)$$

④ 송신 단말에서 전송된 패킷은 전송 지연( $D_T$ ) 후에 수신 단말에 도착한다. 이 때의 지연은 네트워크에서의 전송 지연과 수신 단말의 지터 버퍼 지연을 포함한다.

⑤ 수신 단말의 코덱에 입력된 수신 패킷 중에서 2 번째 프레임의 내용인  $[R(1)|P(2)]$  패킷이 손실된 경우를 나타낸다.

⑥ 2 번째 프레임  $[R(1)|P(2)]$  패킷이 손실되었을 경우에 3 번째 프레임  $[R(2)|P(3)]$  패킷 중에서 R(2) 잉여 정보를 사용해서 복원을 수행한다. 즉, 2 번째 프레임을 복원하기 위해서 3 번째 프레임 패킷의 내용을 사용하기 때문에 3 번째 패킷이 수신된 이후에 2 번째 패킷 복원을 시작해야 한다. 따라서 매체 종속 FEC 기법의 사용으로 인해 한 프레임의 추가 지연이 발생하는데

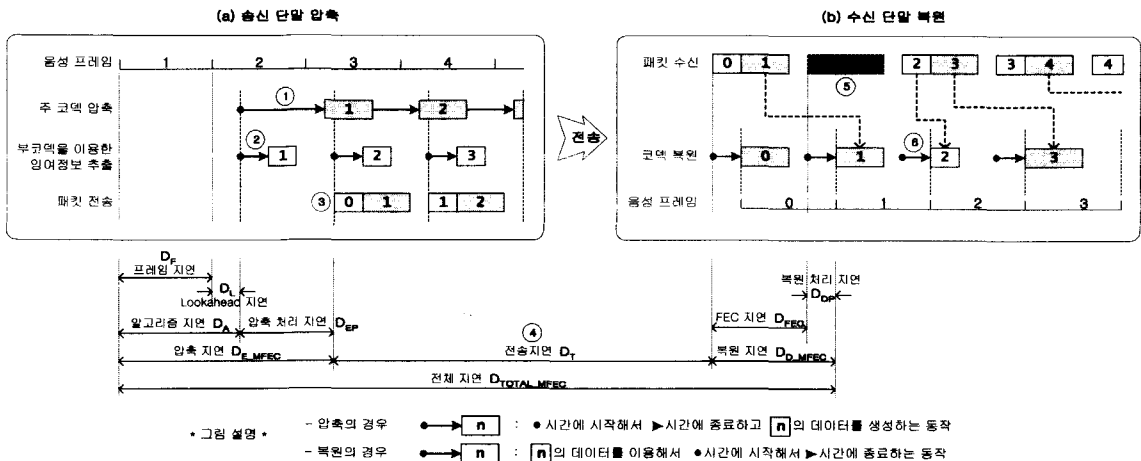


그림 3. 매체 종속 FEC 기법의 타이밍도  
 Fig. 3. Timing diagram of media-specific FEC technique.

이를 FEC 지연(DFEC)이라 한다. 이 때 복원 지연은 식 2와 같이 나타난다. 여기서 DDP는 수신 단말이 패킷 복원을 시작해서 종료하는데 까지 소요되는 처리지연을 말한다.

$$D_{D\_MFEC} = D_{FEC} + D_{DP} \quad (2)$$

위와 같은 전체 흐름을 볼 때 송신/수신 단말 간 전체 지연은 식 (3)과 같이 나타난다.

$$D_{TOTAL\_MFEC} = D_{E\_MFEC} + D_T + D_{D\_MFEC} \quad (3)$$

표 1. 매체 종속 FEC 기법의 지연 요소  
Table 1. Delay source of media-specific FEC technique.

코덱	D <sub>x</sub>			D <sub>T</sub>	D <sub>D</sub>	
	D <sub>F</sub> (ms)	D <sub>L</sub> (ms)	D <sub>EP</sub> (ms)		D <sub>FEC</sub> (ms)	D <sub>DP</sub>
G.723.1	30	7.5	≈ 30	D <sub>T</sub>	30	D <sub>DP</sub>
G.729	10	5	≈ 10	D <sub>T</sub>	10	D <sub>DP</sub>

VoIP에서 가장 많이 사용하는 ITU-T G.723.1과 G.729 코덱에 대한 매체 종속 FEC 기법의 지연 요소는 <표 1>과 같다. 표에서 D<sub>T</sub>와 D<sub>DP</sub>는 본 논문에서 제시하는 방법과 독립적이며 구현에 따른 변수이므로 별도의 수치를 제시하지 않는다. 이를 통해, 매체 종속 FEC 기법은 언어 정보를 이용한 손실 패킷 복원을 통해 음성 품질 개선 효과는 있지만 부수적으로 발생하는 추가 지연으로 인하여 음성 품질 개선 효과를 상당부분 상쇄시킨다는 것을 알 수 있다.

### III. 저지연 LSF FEC 기법

음성 신호를 압축/복원하는 코덱은 waveform 코딩 방식, vocoder 방식, 그리고 이 둘을 합쳐서 사용하는 hybrid 코딩 방식이 있으며, hybrid 코딩 방식이 저전송률(4kbps~8kbps) 대역에서는 가장 좋은 복원 음성 품질을 나타내는 것으로 알려져 있다. VoIP 표준으로 선정된 ITU-T G.723.1, G.729 코덱도 역시 CELP 계열의 hybrid 코딩 기법을 사용하고 있다.

일반적인 hybrid 코딩 기법은 <그림 4>와 같이 전처리, LPC 추출, 피치 추출, residual 코드북 추출의 과정을 수행하고 각각의 과정에서 추출된 계수를 패킷으로 만들어 전송하는 기본 구조를 따른다. LPC 추출 부분

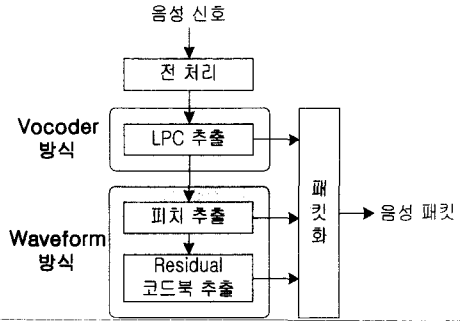


그림 4. Hybrid CELP 코덱의 압축 흐름도  
Fig. 4. Compression flow of hybrid CELP codec.

은 사람의 발성 모델을 기본으로 한 vocoder 방식의 가장 대표적인 예이다. LPC 계수는 일명 formant라고 부르며 사람의 성대에 의한 음성 스펙트럼 특유의 포락선을 모델링 하는 부분으로, 입력 신호의 의미 전달에 가장 큰 부분을 차지하며 전송특성의 개선을 위해서 LSF(Line Spectral Frequency)로 변환해서 사용한다. 그 이후의 피치, residual 코드북 추출은 추출된 LPC를 기반으로 waveform 방식인 perceptual analysis-by-synthesis 방법을 사용해서 후보 파라미터를 모두 적용하고 그 중에서 최적의 파라미터를 검색하는 과정을 수행하며 동일 동작의 반복에 따른 많은 계산량이 소요된다. 따라서 formant를 모델링 하는 LPC 계수가 압축 과정의 가장 선두에서 수행되고, 계산량이 나머지 동작에 비해 적기 때문에 압축 과정의 초기 부분에 추출됨을 짐작할 수 있다. 실제로, [6,7]에 따르면 신호처리 프로세서를 이용한 실제적인 전처리와 LPC 해석 및 LSF 변환에 소요되는 시간 비율은 표 2와 같다.

표 2. LSF 계수 추출 소요 시간 비율  
Table 2. Time rate to compute LSF coefficient.

코덱	G.723.1 (6.3 kbps)	G.723.1 (5.3 kbps)	G.729	G.729A
프레임 크기	30 ms	30 ms	10 ms	10 ms
점유율	8 %	9.3 %	22 %	34.8 %

<표 2>를 보면 전 처리와 LPC 해석에 소요되는 시간 비율은 G.723.1은 10% 이내, G.729는 25% 이내이고, G.729의 계산량을 줄인 버전인 G.729A는 약 35% 정도의 시간이 소요된다. 따라서 VoIP에서는 전체 압축 계산 과정 중에서 적어도 35% 이전에는 LSF 성분의 추



이 때의 복원 지연 ( $D_D$ )은 식 5와 같이 나타난다.

$$D_{D\_LFEC} = D_{DP} \quad (5)$$

위와 같은 전체 흐름을 볼 때 송신/수신 단말 간 전체 지연은 식 (6)과 같이 나타난다.

$$D_{TOTAL\_LFEC} = D_{E\_LFEC} + D_T + D_{D\_LFEC} \quad (6)$$

표 3. Low-delay LSF FEC 기법의 지연 요소

Table 3. Delay source of low-delay LSF FEC technique.

코덱	$D_E$				$D_T$	$D_D$
	$D_F$ (ms)	$D_L$ (ms)	$D_{EF}$ (ms)	$D_{LSF}$ (ms)		
G.723.1	30	7.5	≈ 30	3 이하	$D_T$	$D_{DP}$
G.729	10	5	≈ 10	3.5 이하	$D_T$	$D_{DP}$

표 4. 두 기법의 지연 비교

Table 4. Comparison of delays caused by two technique.

코덱	$D_x$	$D_{TOTAL\_MFEC} - D_{TOTAL\_LFEC}$	지연 차이(ms)
G.723.1		$D_{FEC} - D_{LSF}$	27 이상
G.729		$D_{FEC} - D_{LSF}$	6.5 이상

<표 2>를 참조한 저지연 LSF FEC 기법의 지연 요소는 <표 3>과 같으며 두 기법의 지연 요인을 비교한 결과는 <표 4>와 같다. 저지연 LSF FEC 기법을 사용할 경우 매체 중속 FEC 기법에 비해 G.723.1 코덱은 최소 27ms 이상, G.729 코덱은 6.5 ms 이상의 지연을 단축시키는 효과가 있음을 알 수 있다.

그리고 매체 중속 FEC 기법은 잉여 정보를 생성하기 위해서 코덱과는 별도의 파라미터 추출로 인한 추가 계산량이 요구되지만 저지연 LSF FEC 기법은 코덱에서 생성하는 LSF 정보를 사용하기 때문에 계산량의 증가가 거의 없다는 장점도 있다.

#### IV. 성능 평가

지금까지는 본 논문에서 제시하는 기법의 저지연 특성에 대해서 설명했다. 본 장에서는 저지연 LSF FEC 기법이 적용된 경우와 FEC를 적용하지 않은 경우를 비교해서 저지연 LSF FEC 기법의 음질 개선 효과를

검증할 시뮬레이션을 통한 평가 방법을 설명한다. 성능 평가 시뮬레이션 흐름은 <그림 7>과 같다.

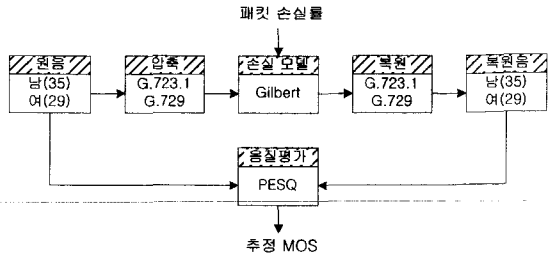


그림 7. 성능 평가 흐름도

Fig. 7. Flow chart of performance evaluation.

평가 환경 요약은 다음과 같다.

1. 음성 데이터 : 64개 음성 신호
2. 코덱 : 6.3kbps G.723.1, 8kbps G.729
3. 성능 비교 : No FEC vs LSF FEC
4. 손실 모델 : Gilbert model
5. 음질 평가 : PESQ
6. 반복 : 10회 이상

#### 1. 음성 데이터

현실적인 실험조건을 위해서 실제 통신 선로 상의 음성 신호를 저장한 음성 데이터베이스를 사용했다. 데이터는 16 비트, 모노, 8KHz 샘플링 된 남성 화자 35명, 여성 화자 29명으로 구성되어 있으며 약 3~7초의 길이를 가진다. 내용은 라디오 교통 정보, 라디오 날씨, 라디오 뉴스, 라디오 사설, 핸드폰 통화 내용 등의 내용으로 구성되어 있다. 실질적인 통신 선로 상의 음성 신호를 저장한 데이터이기 때문에 잡음이 많은 특성을 가진다.

#### 2. 코덱

VoIP에서 가장 널리 사용되는 ITU-T G.723.1, G.729 코덱을 사용하였다. G.723.1 코덱은 6.3Kbps/ 5.3Kbps 두 개의 코덱 중에서 6.3kbps의 전송률을 갖는 MP-MLQ 코덱을 사용했으며, G.729는 계산량을 줄이지 않은 원래 버전을 사용했다. 두 코덱 모두 규격에서 권고하는 오류 은닉 메커니즘을 적용하였다.

#### 3. 성능 비교

VoIP에서 가장 많이 사용하는 G.723.1, G.729 코덱은 각각 30ms, 10ms의 다른 프레임 크기를 가지고 있어서

G.729를 주 코덱으로, G.723.1을 부 코덱으로 한 이중 코덱(Dual Codec) 기법을 사용할 수 없는 환경에 있다. 따라서 두 방식에 대한 성능 비교는 불가능하며 본 논문에서는 FEC를 적용하지 않고 단순히 오류 은닉 메커니즘 만을 적용한 경우와 본 논문에서 제시한 저지연 LSF 기법에 대한 비교를 수행하였다.

4. 손실 모델

인터넷에서는 상황에 따라 다르지만 약간의 손실이 따르는 것으로 알려져 있으며, 손실이 burst 특성을 지니고 있기 때문에 uniform 모델 보다는 연속 손실의 확률이 높은 저 차의 Markov process로 모델링이 가능하다.<sup>[8,9]</sup> 그 중에서도 2차의 Markov 모델인 <그림 8>의 Gilbert 모델이 가장 근사한 모델로 알려져 있다.<sup>[10]</sup>

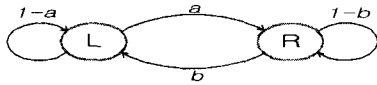


그림 8. Gilbert 손실 모델  
Fig. 8. Gilbert loss model

L 상태가 손실이 발생하지 않은 상태이고 R 상태가 손실이 발생한 상태라고 하고 a를 L 상태에서 R 상태로 천이할 확률, b를 R 상태에서 L 상태로 천이할 확률이라고 하면, 임의의 패킷이 손실될 확률은 R에 머무르는 평균 시간을 R에 머무르는 시간과 L에 머무르는 시간의 합으로 나누어 식 7과 같이 구할 수 있다.

$$P(Loss) = \frac{a}{a+b} \tag{7}$$

그리고 평균 패킷 손실 길이는 R 상태에 머무르는 평균 시간이며 식 8과 같이 구할 수 있다.

$$E(LossBurst) = \frac{1}{b} \tag{8}$$

본 시뮬레이션에서는 b 값을 0.9로 고정하고 손실률(P)은 0%~10%를 사용했다.

5. 음질 평가

일반적으로 음성의 품질을 측정하기 위해서 주관적(objective) 음질 측정법인 MOS (mean opinion score)가 주로 사용되고 있으나 근래에는 시간과 비용의 한계로 인해 MOS를 대체할 만한 객관적(subjective) 혹은 인지적(perceptual) 음질 평가 알고리즘이 많이 제시

되고 있다. 본 논문에서는 각 손실률별 음질을 평가하기 위해서 [11,12]에서 제안한 MOS와 상관도가 90% 이상인 PESQ(perceptual evaluation of speech quality) 알고리즘을 사용했다.

6. 반복

시뮬레이션 결과의 랜덤성을 제거하기 위하여 <그림 7>과 같은 시뮬레이션을 10회 반복하여 수행하였다.

V. 시뮬레이션 결과

본 장에서는 이미 설명한 성능 평가 기법을 통한 시뮬레이션 결과를 제시하고 평가한다.

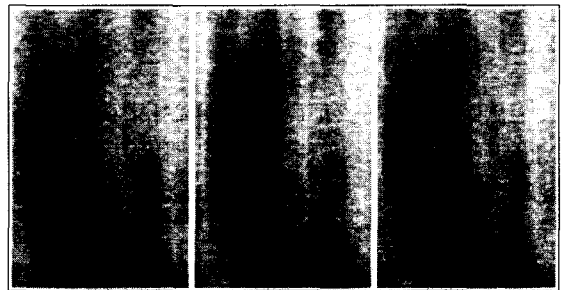


그림 9. 음성 신호 복원 스펙트로그램(G.729)  
(a) 무 손실 환경  
(b) 10% 손실 환경(No FEC)  
(c) 10% 손실 환경(저지연 LSF FEC)

Fig. 9. Spectrogram of decompression speech signal(G.729).  
(a) No lossy environment  
(b) 10% lossy environment(without FEC)  
(c) 10% lossy environment(low-delay LSF FEC).

<그림 9>는 저지연 LSF FEC 기법이 적용된 경우와 그렇지 않은 경우에 대한 손실 복구를 보기 위한 스펙트로그램이다. (a)는 무 손실 환경에서의 G.729 압축/복원 후의 신호, (b)는 10% 손실 환경에서 G.729 코덱 자체의 오류 은닉 메커니즘만이 적용된 신호, 그리고 (c) 그림은 본 논문에서 제안한 저지연 LSF FEC 기법을 사용한 경우의 신호이다. 그림을 비교하면 단순 오류 은닉 알고리즘에서 복원하지 못한 formant를 제안된 기법에서는 많은 부분 복원해 낸 것을 볼 수 있다.

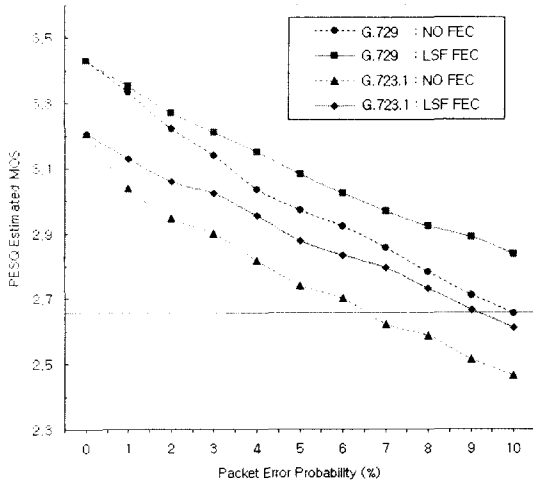


그림 10. G.723.1, G.729 시뮬레이션 결과  
Fig. 10. Simulation results for G.723.1, G.729.

<그림 10>은 G.723.1, G.729 코덱에서의 PESQ 추정 MOS를 시뮬레이션 한 결과이다. 결과를 보면, 두 코덱 모두 손실률이 증가하면 MOS가 감소하는 특성을 가지고 있으며 손실률이 10%가 되면 0.8 가까이 감소한다. 하지만 저지연 LSF FEC 기법을 사용한 경우에는 전체적인 음질 개선 효과를 가지며 최대 10% 손실의 경우에는 MOS 0.2 이상의 음질 개선 효과가 있다. G.723.1의 경우에는 10% 손실 시에 저지연 LSF FEC를 적용하지 않은 G.729와 거의 동일한 음질을 보이는 것으로 나타났다.

<표 5>는 저지연 LSF FEC와 NO FEC 환경에서의 현실적인 손실률인 5%를 가정하고 전체적인 성능 비교를 표로 나타낸 것이다. G.723.1, G.729 코덱 모두 최대 3.5ms의 추가지연으로, 계산량의 증가 없이 MOS 0.1의 통화 품질 개선 효과가 있는 것을 알 수 있다.

표 5. 전체적인 성능 분석 (PER = 5%)  
Table 5. Overall performance analysis (PER = 5%).

코덱	요소	종류	추가 지연	추정 MOS	계산량
G.723.1 6.3kbps	NO FEC		0	2.74	A
	LSF FEC		3ms 이하	2.88	≈A
	효과		3ms 증가	0.14 개선	거의 동일
G.729	NO FEC		0	2.97	B
	LSF FEC		3.5ms 이하	3.08	≈B
	효과		3.5ms 증가	0.11 개선	거의 동일

## VI. 결 론

본 논문에서는 VoIP 음성 패킷 손실 대응 방안으로 제시되고 있는 매체 종속 FEC 기법을 개선하여 패킷 손실 은닉에 사용되는 잉여 정보로 미래 프레임의 LSF 성분을 사용함으로써, 전송 지연을 줄이고 통화 품질을 개선할 수 있는 저지연 LSF FEC 기법을 제안하고 성능을 평가하였다. 성능 평가를 위해서 VoIP에서 사용하는 ITU-T G.723.1, G.729 코덱을 Gilbert 손실 모델에 적용하고, PESQ 음질 측정 방법을 사용하여 각 손실률 별로 MOS를 추정하는 방법을 사용하였다. 본 논문에서 제안한 기법은 기존의 매체 종속 FEC 기법에 비해서 6.5ms~27ms 이상의 지연 감소 효과를 가지고 있는 것으로 나타났으며, FEC를 적용하지 않은 경우와의 복원 음성 품질 비교 시뮬레이션 결과, 5% 정도의 현실적인 손실 환경에서 MOS 0.1 이상의 음질 개선 효과를 보인 것으로 나타났다.

뿐만 아니라, 지금까지 제시된 FEC 기법들은 G.723.1, G.729 보코더의 프레임 크기의 특이성(10, 30ms)으로 인해 VoIP 시스템에 적용하기에 상당히 까다로운 면이 많이 있었으나, 본 논문에서 제안하는 기법은 별도의 계산량 추가 없이 어떠한 프레임 크기에서도 적용이 가능하기 때문에 활용도를 크게 높일 수 있을 것으로 보인다.

## 참 고 문 헌

- [1] B. Duysburgh, S. Vanhastel, B. De Vreese, C. Petrisor, P. Demeester, "On the influence of best-effort network conditions on the perceived speech quality of VoIP connections," Proceedings of IEEE International Conference on Computer Communications and Networks, 2001, pp. 334~339.
- [2] C. Perkins, O. Hodson, V. Hardman, "A Survey of packet loss recovery techniques for Streaming Audio," IEEE Network, Sep./Oct., 1998.
- [3] N. Erdol, C. Castelluccia, A. Zilouchian, "Recovery of missing speech packets using the short-time energy and zero-crossing measure-



- ments," Trans. Speech and Audio Processing, Vol. 1, No. 3, pp. 295~303, Jul. 1993.
- [4] V. Hardman et al., "Reliable audio for use over the Internet," Proc. INET 95, 1995.
- [5] J. -C. Bolot and A. Vega-Garcia, "The case for FEC based error control for packet audio in the Internet," to appear, ACM Multimedia Sys.
- [6] S. M. Lee, S. Park, Y. Jang, "Cost-effective implementation of ITU-T G.723.1 on a DSP chip," Proc. of the 1997 IEEE Int. Sym. on Cons. Elec., pp. 31~34.
- [7] R. Salami, C. Laflamme, B. Bessette, J-P. Adoul, "Description of ITU-T recommendation G.729 ANNEX A: reduced complexity 8 Kbps CS-ACELP codec," ICASSP-97, 1997 IEEE International Conference on ASSP, Vol. 2, pp. 775~778, 1997.
- [8] T. Miyata, H. Fukuda, S. Ono, "Characteristics of successive loss process," Proc. 15th IEEE Int. Conf. on Information Networking, 2001.
- [9] D. Veitch, "Understanding end-to-end Internet traffic dynamics," IEEE Globecom 98, Nov. 1998.
- [10] Dan Rubenstein, Jim Kurose, Don Towsley, "Real-time reliable multicast using proactive forward error correction," UMASS CMPSCI Technical Report pp. 98~99, 1998.
- [11] A. W. Rix, J. G. Beerends, M. P. Hollier, A. P. Hekstra, "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codes," ASSP 2001, Vol 2, pp. 749~752.
- [12] ITU-T Rec. P.862, "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs," Jan, 2002.

## 저 자 소 개



梁海龍(正會員)

1996년 2월 부산대학교 전자공학과 졸업(공학사). 1998년 2월 부산대학교 대학원 전자공학과 졸업(공학석사). 1998년 2월~2001년 5월 현대전자 통신연구소 연구원. 2001년 5월~현재 한국전자통신연구원 부설

국가보안기술연구소 연구원. <주관심분야 : 음성 신호처리, VoIP, 정보보호 등>



李垞勳(正會員)

1992년 2월 고려대학교 전자공학과 졸업 (공학사). 1994년 8월 고려대학교 대학원 전자공학과 졸업 (공학석사). 1998년 8월 고려대학교 대학원 전자공학과 졸업 (공학박사). 1998년

12월~2000년 10월 한국전자통신연구원 선임연구원. 2000년 10월~2001년 11월 (주)텔리언 팀장. 2002년 1월~현재 한국전자통신연구원 부설 국가보안기술연구소 선임연구원. <주관심분야 : 영상처리, 비선형 디지털 신호처리, 정보보호 등>



黃仁浩(正會員)

1980년 2월 한양대학교 전자통신공학과 졸업 (공학사). 1982년 2월 중앙대학교 대학원 전자공학과 졸업 (공학석사). 1992년 2월~1999년 2월 한국과학기술원 대학원 전기전자공학과 졸업 (공학박사). 1986년 2월~

2000년 1월 국방과학연구소 선임연구원. 2000년 2월~현재 한국전자통신연구원 부설 국가보안기술연구소 책임연구원. <주관심분야 : 디지털 신호처리, 통신 시스템, 정보보호 등>