

이중 모드 G.723.1 음성부호화기에 의한 DTMF 톤 전송성능 개선

허호영*, 김희동**, 전준현***, 김형순*

*부산대학교 전자공학과, **한국외국어대학교 정보통신공학과, ***한국통신 멀티미디어 연구소

An Improvement of DTMF Tone Transmission Based on Dual Mode G.723.1 Vocoder

Ho Young Hur*, Hee Dong Kim**, Jun Hyun Jeon***, Hyung Soon Kim*

*Dept. of Electronics Eng., Pusan National Univ., ** Dept. of Information and Comm.Eng.,

Hankuk University of Foreign Studies, ***Korea Telecom Multimedia Tech. Research Lab.

E-main:{hyher, kimhs}@hyowon.cc.pusan.ac.kr, kimhd@ice.hufs.ac.kr

요약

G.723.1은 부호화 방식은 제한된 계산량으로 낮은 전송율에서 음성을 가장 잘 표현할 수 있도록 최적화되어 있어서, 음성주파수 대역에 있는 DTMF 톤의 경우 왜곡이 발생되어 전송성능이 떨어지는 문제점이 있다. 본 논문에서는 DTMF 톤의 투명한 전송을 위해 LSP 계수를 이용한 톤 신호 검출에 기반을 둔 음성모드와 톤모드의 이중모드를 가지는 변형된 G.723.1 음성부호화 방식을 제안한다. 제안된 방식에서 음성모드 부호화기는 기존의 방식과 동일하며, 톤모드의 경우 부호화 단계에서 spectral smoothing 및 피치주기 검출 방식 등을 수정함으로써 수신단의 변경 없이 DTMF 톤의 전송왜곡을 개선시킨다. 본 논문에서는 컴퓨터 모의실험을 통해 제안된 방법이 DTMF 전송성능을 개선시켰음을 확인 하였다.

1. 서론

낮은 전송속도에서 toll quality의 음질을 갖는 G.723.1 부호화기는 최근 VoIP(Voice over IP)에서의 음성부호화 방식으로 각광을 받고 있다. 인터넷폰 등 일반 음성통신 분야에 G.723.1 부호화 방식을 적용하는 데에는 패킷의 오버헤드 양, 패킷의 손실에 대한 견인성 등이 고려되었으나, DTMF(Dual-Tone Multi-Frequency)톤의 투명한 전송을 보장하지 못한다는 결점이 있다.

VoIP(Voice over IP)에서는 이 문제를 해결하기 위하여 DTMF 톤을 별도의 데이터 채널을 통해 전송하는 out-of-band 신호방식을 제안하고 있다. 그러나, 이 out-of-band 신호방식은 수신측에도 DTMF 발생기를 실장하여야 하며, 제어구조가 복잡해지는 단점이 있다. 본 논문에서는 G.723.1 부호화 전단계 또는 그 내부에서 LSP 계수를 이용하여 톤 신호 여부를 검출한 다음, 톤 신호에 대해서는 spectral smoothing 및 피치주기 검출 방식 등을 바꾸어 줌

으로써 톤의 왜곡을 최소화 시키는 톤 모드 부호화기를 동작시키고, 음성신호에 대해서는 기존의 부호화기를 그대로 사용하는 이중 모드 부호화 방식을 제안한다. 제안된 방식을 이용하면 수신단의 변경 없이 DTMF 톤을 in-band로 보낼 수 있어서 전송한 문제를 해결할 수 있다. 본 논문에서는 컴퓨터 모의실험을 통해 제안된 방법이 DTMF 전송성능을 개선시켰음을 확인 하였다.

2. G.723.1 음성 부호화기의 개요

G.723.1은 DSP Group의 TrueSpeech를 기반으로 ITU-T에 의해 표준화된 음성 부호화기이다[1]. 이 표준안은 마니오펜의 상용화에 따른 음성부호화기의 표준화 요구에 의해 ITU-T에서 표준화를 진행하여 1995년 말에 완성되었다. 5.3 및 6.3Kbps의 두 가지 전송율로 동작되는 이 부호화 방식은 음질이 우수하며, 1997년 말 VoIP 포럼에서는 G.723.1을 표준안으로 채택하였다.

G.723.1 부호화기의 알고리즘은 기본적으로 합성에 의한 분석법에 기반을 두고 있으며, 240 샘플(30msec)단위로 프레임(frame)시리를 하는데, 한 프레임은 60 샘플씩을 가지는 부프레임(subframe)으로 나뉘어지 각각의 부프레임에 대해 선형예측 분석을 수행한다. 그리고, 선형예측 분석시에 ill-condition이 되는 것을 방지하기 위해 spectral smoothing, bandwidth expansion 및 white noise correction과 같은 방법을 사용한다. 예 프레임마다 마지막 부프레임의 선형예측 계수는 LSP 계수로 바뀌어지고, 이미 구해져 있는 음성의 LSP 계수 DC 값과의 차이값 PSVQ(predictive split vector quantization)로 양자화 한다. 두 개의 부프레임(120 샘플)마다, 각각가중 필터를 통과한 음성을 이용하여 open-loop 피치를 찾으며, 이때 피치 구간은 18부터 145 사이의 값에서 탐색하게 된다. 추정된 open-loop 피치와 실제 합성필터의 입력스 응답으로 closed-loop 피치예측기를 구성하고 피치는 open-loop 피치주위의 몇 샘플에서 찾게

된다. Close-loop 피치예측 후 잔류신호는 5.3kbps 모드에서는 ACELP 방식, 6.3kbps 모드에서는 MP-MLQ 방식을 사용하여 압축을 하는데, 두 방식 모두 복수의 펄스들에 의해 복호화됐을 때 원음과 평균자승오차를 최소화 하도록 잔류신호를 모델링한다.

3. G.723.1 부호화기에서의 DTMF tone 전송시 문제점 검토

G.723.1은 음성의 고박율 압축을 위해 제안된 것이기 때문에, DTMF 톤이 입력될 경우 peak 주파수의 변경, noise 발생, 한쪽 톤이 다른쪽보다 커지는 twist 등의 왜곡이 발생하여 수신단에서의 검출이 어려워지는 경우가 많이 발생한다[3]. 그림 1(a)는 G.723.1 부호화기에 입력되는 DTMF 신호이고, 그림 1(b)는 복호화되어 나온 톤의 스펙트럼이다. 그림 1(b)를 보면 그림 1(a)에 비해서 고주파 톤이 상대적으로 더 커져 있고, 전반적으로 대역폭이 넓어지며, 두 peak 사이에 또 다른 peak가 생겨 수신단의 DTMF 검출기에서 오류를 유발할 가능성이 거진다.

DTMF 전송왜곡의 문제를 야기시키는 요인들은 다음과 같다[2]

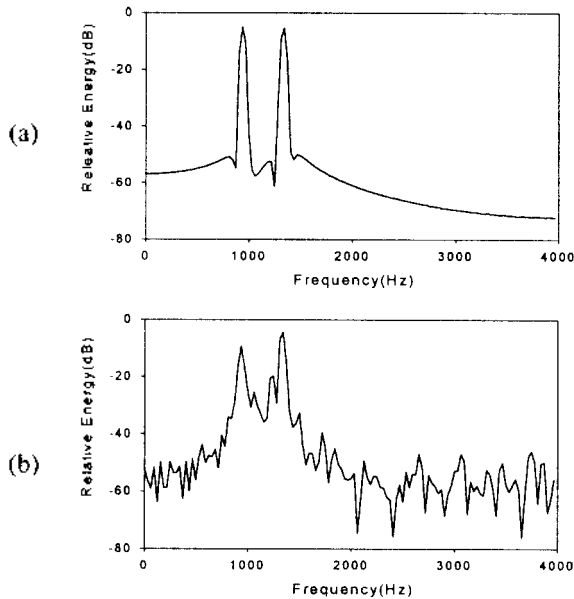


그림 1. G.723.1 부호화기를 통과 했을 때의 스펙트럼 왜곡을 나타내는 예(DTMF 0의 경우)

- (a) 부호화기를 통과하기 전의 스펙트럼
- (b) 부호화기를 통과한 다음의 스펙트럼

3.1 Spectral Smoothing Technique(SST)

Spectral Smoothing Technique은 선형예측 계수를 구할 때 사용하는 자기상관계수에 Gaussian 창을 곱해주는 것인데, 그 결과로 spectral peak가 완만해지고 대역폭이 넓어지며, 선형예측 필터의 pole은 단위위에서 멀어지게 된다[4]. DTMF tone의 경우 SST를 제외시킴으로써 기존의 SST를

사용하여 생기는 spectral envelope의 왜곡을 완화시킬 수 있다.

3.2 Pitch Detection

G.723.1 부호화기는 음성신호의 준주기적 특성을 이용하기 위해 피치를 검출 하는데, DTMF 톤에서 가장 낮은 주파수인 697Hz도 음성의 가장 높은 피치 주파수보다 훨씬 높기 때문에, 잘못 검출된 피치 주파수에 의해 long-term prediction이 수행된다. 기존의 피치주기 검색법은 DTMF 톤의 부호화 품질에 악영향을 미치게 된다.

3.3 LSP Interpolation

G.723.1 부호화기는 LSP를 각 프레임의 네번째 부프레임에 대해서만 구하고 나머지 부프레임은 선형보간해서 사용하는데, 음성의 경우에는 음질지하가 크지 않다. 그러나 DTMF 톤의 경우에는 시작과 끝부분에서 심한 왜곡이 생기며, 매우 짧은 DTMF 톤은 검출 못하게 되는 경우가 많다.

3.4 LSP Quantization

LSP를 양자화할 때, 음성 LSP의 DC 값과 입력신호에서 구한 LSP와의 차이를 PSVQ하게 된다. 이때 LSP의 DC 값은 음성에서 구한 것이기 때문에 DTMF 톤에 적용시킬 경우 양자화 오차가 심하게 된다. DTMF 톤의 DC 값을 사용하면 더 나은 spectral envelope 모델링이 가능하지만, 그 DC 값을 복호화기에서도 가지고 있어야 한다. 본 연구는 현실적인 적용을 고려하여, 복호화기는 변경하지 않는 전제를 두고 있으므로 이 부분은 수정하지 않았다.

3.5 LPC Residual 모델링

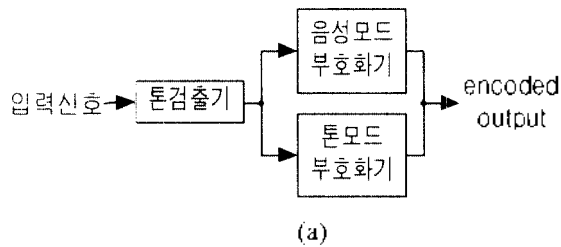
비주기적인 여기신호를 모델링할 때 평균자승오차를 최소화하는 방식으로 구성되어 있기 때문에 DTMF 톤의 경우 대부분 잘못된 모델링을 한다. 즉, DTMF 톤의 선형예측 오차의 스펙트럼에는 두개의 톤 성분이 남아 있지만 G.723.1 부호화기에서 다시 모델링한 여기신호의 스펙트럼은 거의 백색 잡음으로 모델링이 된다. 이러한 여기신호의 모델링이 가장 큰 왜곡을 일으키지만, 여기신호 생성부의 변형은 음성의 특성에 많은 영향을 미치는 부분으로, 복호부분에도 변경이 필요하기때문에 이 부분에 대해서는 고려하지 않았다.

4. DTMF 톤 전송의 개선방법 및 실험

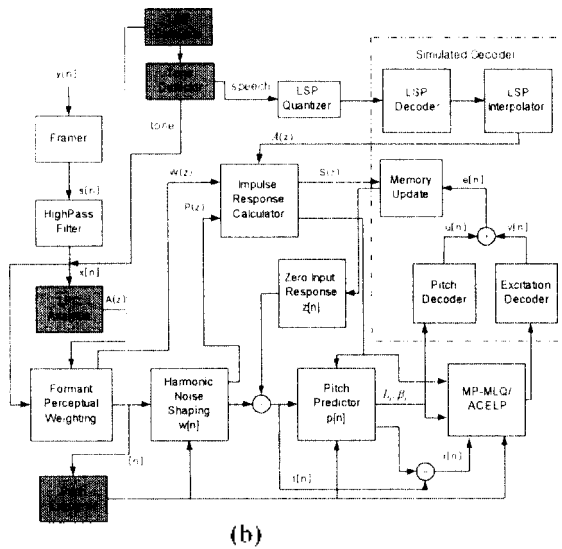
4.1 Dual mode G.723.1 부호화기의 구성

DTMF 톤을 G.723.1 부호화기로 전송시키기 위한 제안된 시스템은 그림 2와 같다. 그림 2(a)는 제안된 이중모드를 가지는 G.723.1 부호화기의 전체 구성도이며, 입력신호가 들어오면 톤검출기에서 톤을 검출하여 음성이면 음성모드로 동작하고 톤이면 톤모드로 동작을 하게 된다. 이때, 음성모드 부호화기는 기존의 G.723.1 부호화기를 그대로 사

용한다. 그리고, 톤검출기는 외부에 별도의 모듈형태로 구성할 수도 있지만, G.723.1 부호화 과정 중 LSP 계수를 구하는 과정에서 검출하도록 구성하였다. 그림 2(b)는 톤모드의 부호화기의 구조이다. 그림 2(b)에서 톤검출기는 부호화 과정에서 구해지는 LSP 값을 이용하여 톤 검출을 하게 된다. 그리고 톤모드의 부호화기에서 LPC analysis, Pitch estimator 블록은 3 상에서 논의된 바에 근거하여 DTMF 톤의 왜곡이 완화되도록 변형시켰다.



(a)



(b)

그림 2. Dual mode G.723.1 부호화기의 구조
(a) Dual mode G.723.1 부호화기의 기본 구성도
(b) 톤모드 부호화기 구조

4.2 LSP 값을 이용한 톤 검출법[5][6]

부호화기에서 LSP 값을 받아 LSP(0)값을 600Hz와 비교한 후 LSP(0)값이 600Hz보다 작으면 음성신호로 결정한다. 만일 600Hz보다 크다면 LSP 값의 간격이 60Hz 이하인 쌍(Pair)의 계수를 계산한다. 만일 그 계수가 2개가 아니면 음성신호로 판단하게 되고, 만일 쌍의 계수가 2개이면 DTMF 톤으로 판단을 내린다. 그런 다음 저주파값과 고주파값을 두개의 쌍으로부터 구할 수 있다.

4.3 톤모드의 부호화기

3 상에서 논의된 문제점들 중 복호화기의 변형없이 부호화기만 수정함으로써 개선시킬 수 있는 항목은 SST와 피치 탐색, LSP interpolation에 의한 문제점이다.

SST에 의한 문제는 DTMF 톤이 검출된 경우 SST를 제외시킴으로써 해결할 수 있다. 그 대신 이와 별도의 Bandwidth Expansion 과정은 그대로 둠으로써 DTMF 톤의

신형에 즉 분석시에 유발될 수 있는 ill-condition 문제를 피하도록 하였다. 그리고 open-loop 피치 탐색기를 변경시켜 톤의 경우 자기상관함수값의 비교시에 새로운 최대값이 이전의 최대값보다 1.25dB 정도 작더라도 갱신 시킴으로써, 될 수 있으면 큰 피치 주기가 선택되도록 하였다. 마지막으로 LSP의 interpolation에 의한 문제를 해결하기 위해 입력 DTMF 톤의 길이를 두 프레임 이상 되도록 하였다.

4.4 실험 및 검토

그림 3에 기존의 G.723.1 부호화가 및 제안된 방식에 의해 DTMF 톤을 압축/복원한 뒤의 스펙트럼 예가 나타나 있다. 톤모드의 부호화기를 통과한 DTMF 톤(그림 3의 점선)과 기존의 G.723.1 부호화기를 통과한 DTMF 톤(그림 3의 실선)의 스펙트럼을 비교하였을 때, 톤모드의 부호화기를 통과한 DTMF 톤이 기존의 방법에 의한 DTMF 톤보다 톤 사이에 존재하던 작은 peak 성분들이 감소되었고, twist에 의한 왜곡도 완화되었음을 확인할 수 있다.

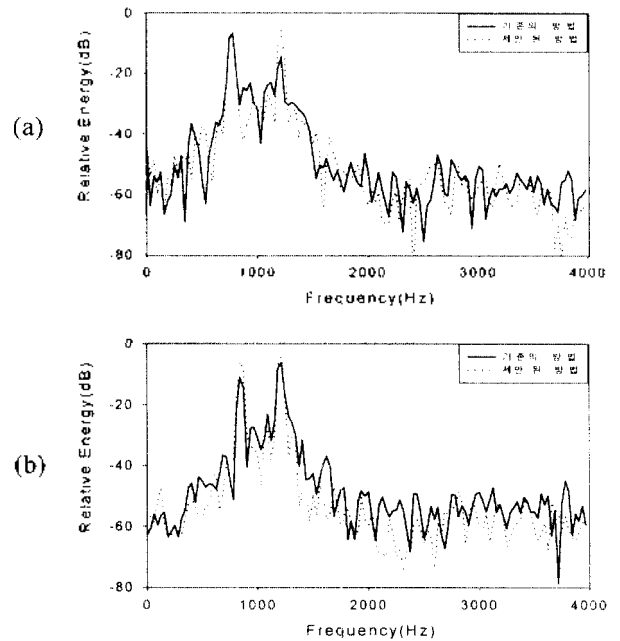


그림 3. 제안된 방식에 의해 DTMF 검출 성능이 개선됨을 보여주는 spectrum의 예

(a) DTMF 톤 4의 경우
(b) DTMF 톤 7의 경우

정확한 성능 평가를 위하여 수신단에서의 DTMF 톤 검출성능을 실험하였다. 그림 4는 컴퓨터 모의실험에 사용한 DTMF 수신기의 구조이다. 입력된 신호는 저주 및 고주 대역의 내역분리여과기를 통과한 다음, 각 신호를 정량화하고 이 결과를 각 톤 여과기로 다시 여과한다. 이때 저주 및 고주 대역의 최대값들을 정해진 임계값과 비교하여, 임계값 이상이면 DTMF 톤으로 판정하고 둘중 하나라도 임계값보다 작으면, DTMF 톤이 아닌 것으로 판정한다. 컴퓨터 모의실험에서는 음성신호와 잡음에 의한 DTMF 톤

loss율과 DTMF 톤이 없을 때 음성과 잡음에 의한 오검출율을 살펴 보았다.

그림 5는 DTMF 전력대 잡음전력이 0dB 일 때 SPEECH/DTMF 전력에 따른 DTMF 톤 Loss율을 나타낸다. 그림에서 선 동그라미는 기존의 방법에 의해 구한 톤의 DTMF 검출 성능을, 삼각형은 톤모드의 부호화기를 통과한 톤의 성능이고, 견성 동그라미는 부호화기를 거치지 않은 원음의 DTMF 검출 성능을 나타낸다.

그림에서 알 수 있듯이 SPEECH/DTMF 전력비가 4dB 이하면 DTMF loss rate가 거의 0%가 되어 원음과 DTMF 검출성능이 비슷해지는 것을 알 수 있다. 그리고, 이때의 임계값은 그림 6에 의해 False AlarmRate가 0%가 되는 150으로 설정하였다.

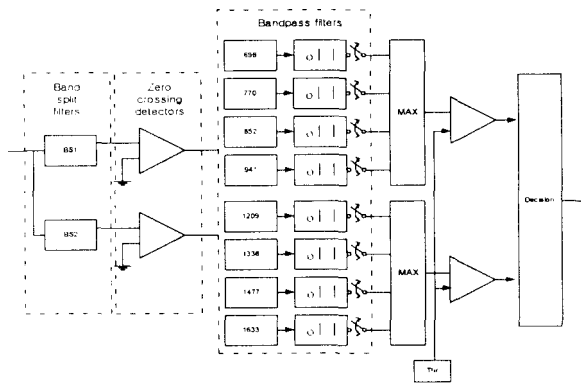


그림 4. 컴퓨터 모의실험시 DTMF 검출기의 구조

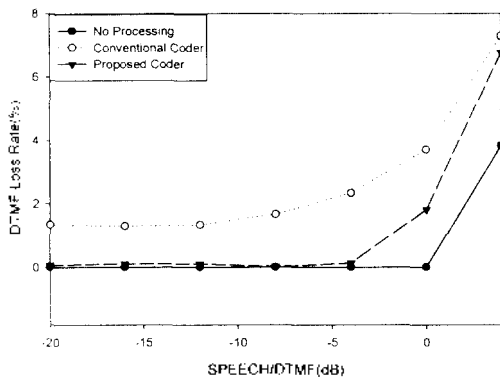


그림 5. DTMF tone Loss 율(threshold = 150)

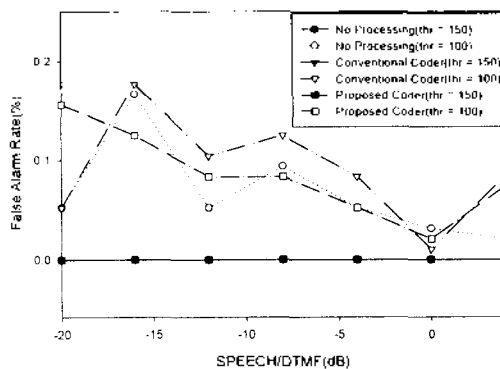


그림 6. False Alarm 율

5. 결론

G.723.1은 부호화 방식은 제한된 계산량으로 낮은 전송율에서 음성을 가장 잘 표현할 수 있도록 최적화되어 있어서, 음성주파수 대역에 있는 DTMF 톤의 경우 왜곡이 발생되어 전송성능이 떨어지는 문제점이 있다. 본 논문에서는 DTMF 톤의 투명한 전송을 위해 LSP 계수를 이용한 톤 신호 검출에 기반을 둔 음성모드와 톤모드의 이중모드를 가지는 변형된 G.723.1 음성부호화 방식을 제안하였다. 제안된 방식은 G.723.1 부호화 전단계 또는 그 내부에서 LSP 계수를 이용하여 톤 신호의 여부를 검출한 다음, 톤 신호에 대해서는 spectral smoothing 및 피치주기 검출 방식 등을 바꾸어 줌으로써 톤의 왜곡을 최소화 시키는 톤 모드 부호화기를 동작시키고, 음성신호에 대해서는 기존의 부호화기를 그대로 사용하게 된다. 수신단에서의 스펙트럼 관찰결과 제안된 방법이 기존의 부호화기보다 왜곡이 감소되며, 컴퓨터 모의실험을 통해 제안된 방법이 DTMF 전송성능을 개선시킴을 확인 하였다.

참고문헌

- [1] ITU Recommendation, "G.723.1 - Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3kbps/s," Mar. 1996.
- [2] 허호영, 김희동, 진준현, 김형순, "G.723.1 음성부호화기에서의 DTMF 톤 전송에 관한 연구," 제 11 회 신호처리학술대회, pp.289-292, 1998년 10월.
- [3] R. V. Cox, "Three New Speech Coder from the ITU Cover a Range of Applications," IEEE Commun. Mag., pp40-47, Sep 1997.
- [4] W. B. Kleijn and K. K. Paliwal, *Speech Coding and Synthesis*, ELSEVIER SCIENCE B V., pp.233-236, 1995.
- [5] 손상문, 박순, "저전송률 음성부호화기의 Dual-Tone MultiFrequency Signaling 검출법," 제 15 회 음성통신 및 신호처리 워크샵 논문집, pp.159-164, 1998년 8월.
- [6] E. F. Tzeng, Rockville, "Dual-Tone MultiFrequency Signaling Transparency for Low-Data-Rate Vocoder," United States Patent No. 5,459,784, 1995.

본 연구는 1998년 한국통신 장기기초 연구과제의 연구비 지원을 통하여 이루어졌습니다.