[5A-96]

Transform 기반 분할대역 광대역 음성부호화기 음질 향상을 위한 시간축 정규화 기법

김경태, 정성교, 이미숙†, 강홍구, 윤대희
연세대학교 전기전자공학과
†ETRI 광대역 통합망 연구단
kktae@mcsp.yonsei.ac.kr

Temporal Normalization Techniques for Transform-type Split-band Wideband Coder

Kyung-Tae Kim, Sung-Kyo Jung, MiSuk Lee†, Hong-Goo Kang, and Dae Hee Youn
Department of Electrical and Electronic Eng., Yonsei University, Korea
†Broadband Convergence Network Division, ETRI

요 약

본 논문에서는 분할대역 방식 광대역(0.5-7 kHz) 음성부호화기를 위한 효율적인 고대역(4-7 kHz) 음성부호화 알고리듬에 대해 연구했다. 분할대역 방식 광대역 음성부호화기는 광대역 음성 신호를 대역통과 필터를 통해 저대역(0-4 kHz)와 고대역(4-7 kHz) 신호로 나누어 각각 독립된 방식으로 부호화한다. 고대역 음성부호화를 위해 Transform 부호화 알고리듬을 사용할 경우, 저전송률 조건에서 충격 신호와 같은 시간축 에너지 변화가 급한 신호에서 귀에 거슬리는 양자화 잡음이 발생한다. Transform 기반 고대역 음성 부호화에서 발생하는 양자화 잡음을 줄이기 위해 본 논문에서 두가지 방법을 제시했다. 첫번째는 시간축 에너지를 직접 계산하여 시간축에서 정규화를 실시하는 "직접 시간축 정규화 방법" 이고, 다른 하나는 주파수 영역에서 선형예측 분석을 실시하여 주파수 영역 단구간 상관도를 제거하는 "주파수영역 선형예측" 방식이다. 실험 결과 제안된 두가지 시간축 정규화 방법을 통해 고대역 신호 부호화 과정에서 발생하는 양자화 잡음을 성공적으로 제거할 수 있었다. 시간축 정규화 방법을 분할대역 광대역 음성부호화에 적용해 본 결과, 20 kbit/s 전송률에서 56 kbit/s ITU-T G.722 보다 나은 음질을 보였다.

Ⅰ. 서 론

IP 네트워크를 통한 패킷통신의 수요급증과 더불어 전 세계적으로 IMT-2000 네트워크와 같은 IP 기반의 무선 통신 시스템이 급속도로 확산되고 있다. 따라서 IP 를 통한 통신은 점차 유/무선 통신 네트워크 모두에 있어 가장 주요한 통신 매체로 진화할 것으로 보인다. 이와같이 유/무선 통신 시스템이 같은 IP 네트워크로 통합됨에 따라 무선 네트워크와 유선 네트워크 간에 발생할 수 있는 IP 패킷의 변환(상호부호화) 문제가 화제로 떠오르고 있다. 그러므로, 새로 개발되는 음성부호화기는 채널 대역을 효율적으로 사용할 수 있을 뿐만 아니라 상호부호화의 필요성을 줄여줄 수 있는 방향으로 설계되어야 한다. 특히, 오랜 시간동안 IP 네트워크에서 사용되고 있는 ITU-T(International Telecommunication Union Telecommunication Standardization Section) G.723.1 [1]이나 G.729[2], G.729 확장 버전들(G.729 annexes)과 쉽게 상호부호화가 가능해야 한다.

ITU-T 는 VoIP 환경을 위한 가변전송률 광대역 음성 부호화기에 대한 표준화 작업을 진행 중이다. 가변전송률 광대역 음성 부호화기는 다양한 채널 환경을 갖는 IP 망에서 채널 상황에 따라 전송률을 변화시켜 효율적인 전송이 가능할 뿐 아니라, 광대역(50-7000 Hz) 음성신호를 처리하므로 고품질의 음성통신에 적합하다는 장점을 가지고 있다. 이를 위해, 현재 두 가지의 접근 방식을 중심으로 표준화 작업이 이루어지고 있다. MSC-VBR(Multirate Source Controlled Variable Bit-Rate)와 EV(Embedded VBR)가 그것이다[3]. 현재로는 이 두 분야 모두 표준화를 위한 ToR(Terms of References) 협의 중에 있다.

최근 가변 대역폭 특성을 갖고 전화대역 부호화기와 연동이 가능한 다양한 가변 전송률 부호화기에 대한 연구가 이루어졌다. 그 중 [4]에서는 입력된 광대역 음성 신호를 QMF (Quadrature Mirror Filter)[5]을 통해 저대역과 고대역으로 나누어 각각을 독립된 방식으로 부호화하는 분할대역(split-band) 방식을 채택했다. 여기서 저대역은 VoIP 시스템에서 많이 사용되고 있는 전화대역 부호화기인 G.723.1 을 통해 부호화하고 저대역의 합성 음성의 품질을 높이기 위해 추가 부호화 모듈을 사용한다. 고대역 신호는 MDCT(Modified Discrete Cosine Trans-form)[6]을 이용한 transform 기반 부호화기를 통해 부호화 된다. 이 분할대역 광대역 음성부호화기는 G.723.1 과 비트 호환이 가능하며 19.4 kbit/s 전송률로 ITU-T 24 kbit/s G.722.1[7] 광대역 음성부호화기와 동등한 음질을 제공한다. 하지만 [4]에서 제안된 접근 방법으로 고대역 음성을 부호화할 경우 onset

본 연구는 한국전자통신연구원(ETRI)의 지원으로 수행되었습니다.