

시간 영역에서 개선된 파라미터 추론을 통한 효율적인 초광대역 확장 시스템 설계

전종근
LIG넥스원

Designing of efficient super-wide bandwidth extension system using enhanced parameter estimation in time domain

Jong-jeon Jeon
LIG Nex1
E-mail : jonggeun.jeon@lignex1.com

요 약

본 논문은 광대역 음성의 음질 향상을 위해 시간 영역에서 인공대역 확장 기술을 사용하여 초광대역 음성신호를 출력하여 사용자에게 개선된 음질의 음성을 제공하는 시스템을 제안한다. 시간 영역에서 소스 필터 모델에 기반하여 광대역 여기신호 및 LSP를 추출하고, 각각의 대역폭 확장 알고리즘을 적용했고, 초광대역 여기신호 및 LSP를 추론하여 초광대역 음성신호를 합성한다. 주관적인 테스트를 통해 광대역 음성신호보다 초광대역 음성신호의 음질을 더 선호하는 결과를 도출하였다.

ABSTRACT

This paper proposes the system that offer super-wideband speech which is made by artificial bandwidth extension technique using wideband speech signal in time-domain. wideband excitation signal and line spectrum pair(LSP) are extracted based on source-filter model in time-domain. Two parameters are extended by each bandwidth extension algorithms, and then, super-wideband speech parameters are estimated. and synthesized. Subjective test shows super-wideband speech is better speech quality than wideband speech signal.

키워드

bandwidth extension, super-wideband, time-domain, source-filter model

1. 서 론

오늘날 음성 통화 통신을 위해 개발되는 음성 코덱들은 광대역 음성신호(wideband speech) 뿐만 아니라, 초광대역 음성신호(super-wideband speech)의 대역폭까지 지원한다. 광대역 음성은 협대역 음성신호에 비해 많은 대역폭을 포함하지만 여전히 사람의 가청주파수 대역에 비해 좁은 대역을 가지고 있기 때문에 여전히 실제 음성에 비해 음성품질이 떨어진다. 반면 초광대역 음성신호는 사람의 가청주파수와 비슷한 대역까지 포함하고 있기 때문에 많은 음성 정보들을 가지고 있어 원음성과 크게 차이가 없을 정도로 좋은 음성 품질을 출력

한다. 하지만 대부분의 모바일 기기나 네트워크 망은 초광대역 음성을 지원하지 못하고 있는 실정이다. 이렇게 제한적인 네트워크 또는 모바일 기기에 서도 초광대역 음성의 통화품질로 제공할 수 있는 기술 중 하나가 바로 인공대역확장(artificial bandwidth extension) 기술이다.[1] 인공대역확장 기술은 추가적인 비트소모 없이 광대역 음성 파라미터들을 초광대역 음성 파라미터로 추론하여 초광대역 음성을 추출할 수 있는데 [2], 본 논문에서는 시간영역에서 초광대역 음성 출력을 위한 인공대역확장 시스템에 대해 설명하고 실험결과를 통해 제안된 시스템의 성능을 평가하였다.

II. 인공대역 확장 시스템 설계

시간영역에서 인공대역 확장 시스템은 소스-필터 모델에 기반하여 각각 광대역 여기신호(wideband excitation signal)와 LSP(Line Spectrum Pair)를 추출하여 각각의 대역 확장 알고리즘에 적용하여 초광대역 여기신호와 LSP를 추출하도록 시스템을 설계하였다.

그림1.은 초광대역 확장 시스템의 설계 구조이다.

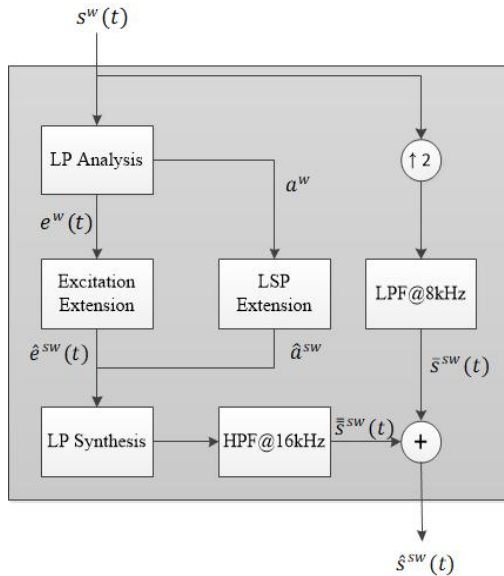


그림 1 초광대역 확장 시스템

광대역 음성 코덱을 통해 디코딩된 광대역 음성 신호 $s^w(t)$ 는 LP분석을 통해 여기신호 $e^w(t)$ 와 LPC 계수 a^w 를 추출한다. 추출된 광대역 여기신호와 광대역 LPC 계수는 각각의 알고리즘을 통해 초광대역 여기신호 $e^{sw}(t)$ 와 LPC 계수 a^{sw} 를 추론한다. 본 논문에서는 초광대역 여기신호 추출을 위해 스펙트럼 시프팅(Spectrum shifting) [1] 방식을 사용하였고, 기존 알고리즘과 비교하여 개선된 초광대역 LSP 파라미터 추론을 위해 선형매핑(Linear mapping)과 코드북 매핑(codebook mapping) [3]을 동시에 사용하여 각 매핑을 통해 얻은 파라미터에 각각 가중치를 주어 연판정(soft-decision) 방식을 사용하였다. 각각 확장된 초광대역 파라미터들은 LP 합성을 통해 추론된 초광대역 음성신호를 생성하고, 8kHz 이상의 고대역 부분만 추출하기 위해 고대역 통과 필터(HPF)를 사용하여 $\bar{s}^{sw}(t)$ 신호를 추출한다. 광대역 입력신호는 업샘플링(Up-sampling)과 저대역 통과 필터(LPF)를 사용하여 광대역 성분만 존재하는 초광대역 음성신호 $\bar{s}^{sw}(t)$ 를 생성한 후, 대역확장을 통해 얻은 $\bar{s}^{sw}(t)$

와 더하여 출력이 되는 초광대역 음성신호 $\hat{s}^{sw}(t)$ 를 생성한다.

그림1은 업샘플링된 광대역 음성 원음의 스펙트로그램(spectrogram)이다. 스펙트로그램에서 확인할 수 있는 것처럼, 7 ~ 14kHz 대역에서 음성 데이터들이 비어있는 것을 확인할 수 있다. 이 광대역 음성을 초광대역 확장 시스템의 입력으로 사용하여 초광대역 확장시스템을 통해 출력된 초광대역 음성신호의 스펙트로그램은 그림2와 같다. 그림2에서 볼 수 있는 것처럼, 인공대역 확장 기술을 통해 7 ~ 14kHz 주파수 대역의 음성 성분이 대역확장 알고리즘을 이용하여 생성된 것을 볼 수 있다.

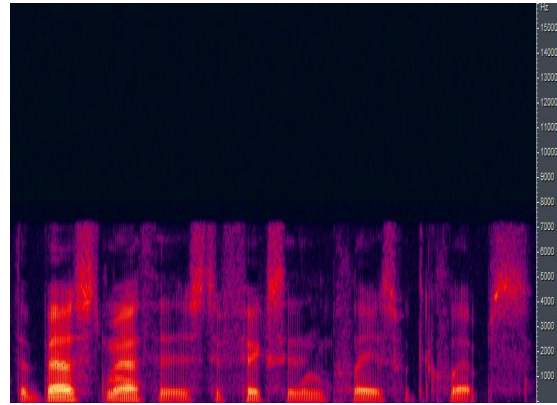


그림 2 광대역 음성 원신호

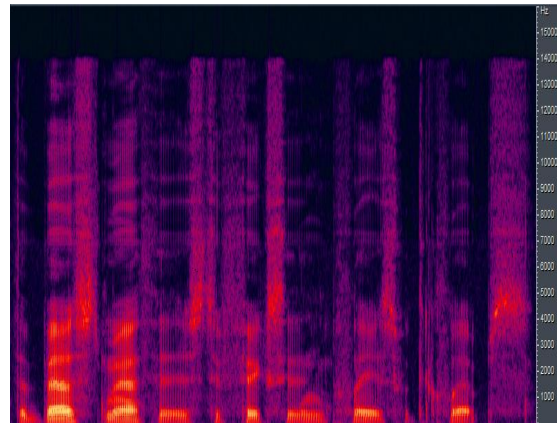


그림 3 초광대역 확장 시스템을 통해 출력된 초광대역 음성 합성신호

III. 실험 결과

실험결과를 도출하기 위해 광대역 입력신호와 초광대역 출력신호를 A/B Test를 수행하였다. 순서에 상관없이 10명의 실험대상자에게 두 음성을 들려주었고 표 1에서 확인할 수 있는 것처럼, 실험 결과 80%가 초광대역 음성의 음질을 더 선호하는 결과를 얻었다.

표 1 선호도 결과

	광대역 음성	초광대역 음성
선호도	20%	80%

IV. 결 론

본 논문에서는 기존 광대역 음성코덱의 변화 없이 보다 나은 음성 품질을 위해 인공대역 확장 시스템을 설계하여 초광대역 음성을 제공하였다. 각각의 대역확장 알고리즘을 통해 초광대역 파라미터를 추론하여 초광대역 음성신호를 출력하였고, 선호도 조사에서도 광대역 원음성보다 합성된 초광대역 음성을 더 선호하는 결과를 얻었다.

References

- [1] A. Fuemmeler, C. Hardie, R. Gardner “Techniques for The Regeneration of Wideband Speech from Narrowband Speech.”, *EURASIP J. Adv. Sig. Proc.* 2001(4), pp. 266-274, 2001.
- [2] Y. H. Lee and S. H. Choi, “Superwideband Bandwidth Extension Using Normalized MDCT Coefficients for Scalable Speech and Audio Coding,” *Hindawi Publishing Corporation Advances in Multimedia vol. 2013*, 2013
- [3] Peter Jax and Peter Vary, “On Artificial Bandwidth Extension of Telephone Speech,” *Signal processing*, Vol. 96, No. 6, pp. 1276-1306, 2006.