

시간 영역 에너지 Companding을 사용한 프리 에코 감소 방법

*김재원 임유진 유정찬 서은미 박호종

광운대학교

*meep7174@naver.com

Pre-Echo Reduction Using Time Domain Energy Companding

*Kim, Jaewon Lim, Yujin Yu, Jeongchan Seo, Eunmi Park, Hochong

Kwangwoon University

요약

본 논문에서는 시간 영역 에너지의 companding을 이용하여 오디오 부호화에서 발생하는 프리 에코를 효과적으로 감소시키는 방법을 제안한다. 일반적으로 오디오 부호화는 블록 단위의 변환 부호화를 사용하므로 과도 구간에서 프리 에코를 발생시킨다. 프리 에코를 줄이기 위한 기존 TNS 방법은 주파수 영역에서 선형 예측 방법을 사용하며, 부가 정보 전송이 필요하고 성능이 낮은 문제점을 가진다. 제안하는 방법은 시간 영역 에너지의 동적 범위를 감소시킨 후 부호화 하고, 복호화 이후에 에너지를 복원하는 과정을 통하여 양자화 오차의 시간 영역 에너지 분포를 조정하여 프리 에코를 감소시킨다. 제안하는 방법이 TNS보다 우수한 프리 에코 감소 성능을 가지는 것을 확인하였다.

1. 서론

변환 부호화(transform coding)를 사용하는 일반적인 오디오 부호화 기술은 정상상태(stationary) 신호의 부호화 이득을 높이기 위해 긴 프레임(frame)을 사용한다. 부호화 과정에서 생긴 양자화 오차는 프레임의 시간 영역에서 균일한 에너지 분포를 가진다. 따라서 과도(transient) 신호를 가지는 프레임에서는 신호 에너지가 작은 구간에서 양자화 오차가 심각한 청취 잡음을 생성한다. 이러한 프리 에코(pre-echo)를 감소시키기 위해 윈도우 스위칭(window switching), temporal noise shaping (TNS), QMF 영역에서의 에너지 companding 방법 등이 개발되었다[1-2].

TNS는 주파수 영역에서의 선형 예측을 사용하여 시간 영역에서 평탄한 신호를 양자화 하고, 복호화 과정에서 시간 축 포락선(temporal envelope)을 복원하여 양자화 오차가 신호와 동일한 시간 축 포락선을 갖도록 한다. 그러나 TNS는 선형 예측 정보를 전송해야 하며 우수한 프리 에코 감소 성능을 제공하지 못한다. 따라서 [2]의 에너지 companding 방법을 기반으로 추가 정보를 사용하지 않고 성능이 우수한 프리 에코 감소 방법을 제안한다.

제안하는 방법은 시간 영역에서 짧은 구간 단위로 계산한 에너지를 companding 하여 신호의 에너지 동적 범위를 제한한다. 부호화 과정에서 양자화 오차가 추가된 신호는 복호화 후에 companding을 통해 기존 신호의 에너지로 복원된다. 이 과정에서 양자화 오차는 신호와 동일한 시간 영역 에너지 분포를 가지며 프리 에코가 감소한다.

제안하는 방법이 TNS보다 우수한 성능을 가지는 것을 복호화기 출력 파형과 스펙트로그램으로 확인하였다. 또한 companding 곡선이 성능에 영향을 주는 것을 확인하였다. 따라서 머신러닝 기반으로 최적의 companding 곡선을 설계하여 성능 향상을 얻을 수 있을 것이다.

2. 제안하는 방법

2.1 시간 영역의 에너지 companding 방법

제안하는 방법은 시간 영역에서 64-샘플 구간 단위로 에너지를 구한다. 각 구간의 에너지는 시간 영역에서 완만한 변화를 가지도록 64-샘플 길이의 신호에 윈도우를 적용하여 계산한다. 식 (1)의 companding 곡선을 통해 에너지를 companding하여 신호의 동적 범위를 제한하고, 식 (1)은 구체적 동작을 결정하는 파라미터 α 를 가진다.

$$y = x^\alpha \quad (1)$$

동적 범위가 감소한 신호는 부호화 과정을 거치며 양자화 오차를 포함한다. 복호화기는 동일한 방법으로 양자화 오차가 포함된 신호의 에너지를 계산하고, companding 곡선의 역을 사용하여 동적 범위가 감소하기 전 에너지를 구하고, 이를 적용한 최종 신호를 복원한다. 이 때 시간 영역에서 균일하게 퍼져있는 양자화 오차에도 동일한 에너지 분포가 적용되어 프리 에코가 감소한다. 단, 부호화 후 에너지는 양자화 오차를 포함하여 원 에너지와 다르게 되고, 따라서 완벽하게 기존 에너지를 복원하지는 못한다.

2.2 Companding에 의한 프리 에코 분석

Companding 곡선의 파라미터 α 를 변경하며 companding 곡선이 프리 에코의 감소 성능에 미치는 영향을 비교한다. 부호화 후 에너지는 양자화 오차를 포함한다. Companding 곡선의 기울기가 큰 경우 양자화 오차가 포함되더라도 복원 값은 원래 입력 값과 오차가 작다. 그러나 기울기가 작은 경우 동일한 양자화 오차가 포함되더라도 복원 값은 더 큰 오차를 가진다. 프리 에코는 에너지가 작은 구간에서 발생하는데, 그림 1의 companding 곡선에서 (b)가 작은 입력에 대해 큰 기울기를 가진

다. 따라서 그림 1의 (b)를 companding 곡선으로 사용한 그림 2의 (c) 신호가 (b) 신호보다 프리 에코를 더 잘 감소시키는 것을 확인할 수 있다. 그러나 에너지가 큰 경우, 그림 1의 (b)의 기울기가 작아 양자화 오차로 인한 복원 성능이 낮은 트레이드오프(Trade-off)가 있다.

그림 3의 (a)는 부호화기 이득으로 식 (1)의 α 에 따라 차이가 있다. 그러나 양자화 오차의 영향으로 복호화기 이득인 그림 3의 (b)는 (a)의 완벽한 역이 되지 못한다. 20 ~ 40 ms 구간에서 그림 3의 (a)와 달리 (b)의 이득은 α 와 상관없이 크기가 비슷하다. 양자화 오차로 인해 큰 이득이 곱해진 그림 2의 (b) 신호는 프리 에코 감소 성능이 낮은 것을 확인할 수 있다. 따라서 복원 성능을 높이기 위해 머신러닝 기반으로 최적의 companding 곡선을 설계하고, 양자화 오차 레벨을 추정하여 에너지 복원 성능을 높여 높은 프리 에코의 감소 성능을 기대할 수 있다.

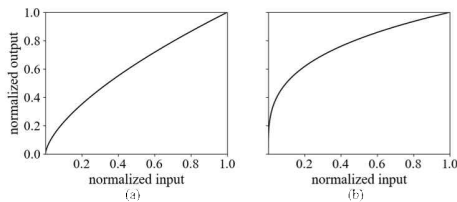


그림 1. Companding 곡선 (a) $\alpha = 0.65$, (b) $\alpha = 0.30$.
Fig. 1. Companding curve (a) $\alpha = 0.65$, (b) $\alpha = 0.30$.

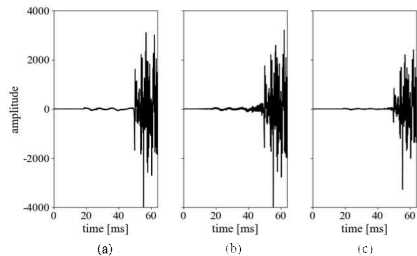


그림 2. Companding 곡선을 사용하여 복원한 신호 (a) 원본, (b) $\alpha = 0.65$, (c) $\alpha = 0.30$.
Fig. 2. Reconstructed signal using companding curve (a) original, (b) $\alpha = 0.65$, (c) $\alpha = 0.30$.

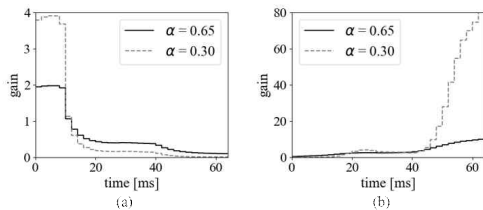


그림 3. 신호에 적용되는 이득 (a) 부호화기, (b) 복호화기.
Fig. 3. Gain applied to signal (a) encoder, (b) decoder.

3. 성능 평가

프리 에코 감소의 성능 평가를 위해 과도 구간의 음성 신호를 사용하였고, 공정한 비교를 위해 모든 프리 에코 감소 방법에서 동일한 unified speech and audio coding (USAC) 부호화 방법을 사용하였다 [3]. 성능 비교는 원본, TNS, 샘플 단위 포락선을 사용하는 방법, 제안하는 방법에 대하여 실시하였다. 샘플 단위 포락선은 Hilbert 변환을 통해 실제 포락선을 계산하고 중앙값 필터(median filter)를 사용하여 저대역 필터링 하여 구하였다. 샘플 단위 포락선 방법은 양자화 오차에 샘플 단위의 실제 포락선을 적용하여 이상적으로 프리 에코를 감소시키지만, 포락선 정보를 전송하기 위해서 많은 비트가 필요하므로 실제 부호화에 적

용하기는 어렵다.

그림 4와 그림 5는 여러 가지 방법으로 프리 에코를 감소시킨 신호의 파형과 스펙트로그램을 보여준다. 제안하는 방법이 TNS 보다 우수한 프리 에코 감소 성능을 보여준다. 샘플 단위의 실제 포락선을 사용한 신호는 균일한 양자화 오차에 샘플 단위로 포락선을 곱하여 프리 에코가 가장 많이 감소 한 것을 확인할 수 있다.

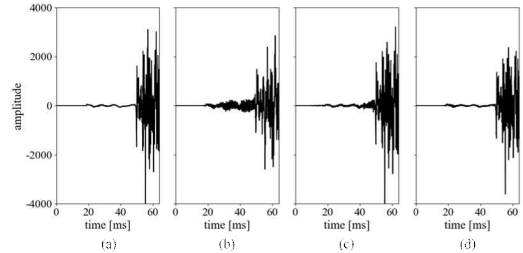


그림 4. 각 프리 에코 감소 방법에 대한 복호화기 출력 신호 (a) 원본, (b) TNS, (c) 제안하는 방법, (d) 샘플 단위 포락선 방법.
Fig. 4. Decoded signals by different methods of pre-echo reduction (a) original, (b) TNS, (c) proposed method, (d) sample-based envelope method.

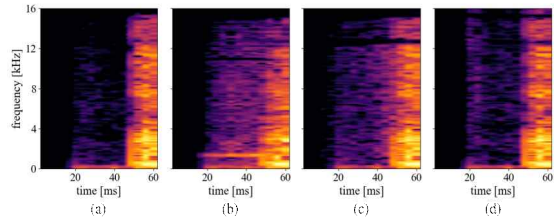


그림 5. 각 프리 에코 감소 방법에 대한 복호화기 출력 스펙트로그램 (a) 원본, (b) TNS, (c) 제안하는 방법, (d) 샘플 단위 포락선 방법.
Fig. 5. Spectrogram by different methods of pre-echo reduction (a) original, (b) TNS, (c) proposed method, (d) sample-based envelope method.

4. 결론

본 논문에서는 시간 영역에서 에너지 companding 방법을 사용하여 프리 에코를 감소시키는 방법을 제안하였다. TNS와 비교한 결과 우수한 감소 성능을 확인하였다. 또한 머신러닝 기반으로 companding curve의 설계와 양자화 오차 레벨 추정을 통해 향상된 성능을 기대할 수 있다.

감사의 글

본 연구는 한국전자통신연구원 연구운영비 지원사업의 일환으로 수행되었음(22ZH1200, 초실감 입체공간 미디어 콘텐츠 원천기술 연구).

참고문헌

- [1] J. Herre and J.D. Johnston, "Enhancing the performance of perceptual audio coders by using temporal noise shaping (TNS)," *Audio Engineering Society Convention 101*, 1996.
- [2] A. Biswas, P. Hedelin, L.F. Villemoes, and V. Melkote, "Temporal Noise Shaping with Companding," in *Proc. INTERSPEECH*, pp. 3548-3552, 2018.
- [3] ISO/IEC 23003-3, "MPEG audio technologies-Part 3: Unified speech and audio coding." 2012.