

# 상호부호화기의 후처리 필터와 인지가중 필터를 대신하는 새로운 필터 설계 및 성능 평가

최진규, 윤성완, 강홍구, 윤대희  
연세대학교 전기전자 공학과 미디어통신 신호처리 연구실

## New filter design to replace the post and perceptual weighting filter of transcoder and performance evaluation

Jin-Kyu Choi, Sung-Wan Yoon, Hong-Goo Kang, Dae-Hee Youn  
Dept. of Electrical & Electronic Eng., Yonsei University  
E-mail : [jinkyu@mcsp.yonsei.ac.kr](mailto:jinkyu@mcsp.yonsei.ac.kr)

### Abstract

*In speech communication systems where two different speech codecs are interoperated, transcoding algorithm is a good approach because of its low complexity and improved synthesized speech quality. This paper proposes an efficient method to further improve the performance of transcoding algorithms as well as to reduce the complexity. In the conventional transcoding algorithms, a post-filter and a perceptual weighting filter should be operated sequentially because both decoding and encoding processes are needed. This results in the redundancy of the processing in terms of complexity and perceptual quality. Using the fact that their filter structures are similar, we replaced the two filters with one. The proposed algorithm requires 72.8% lower complexity than the conventional transcoding algorithm when we compare only the complexity of the filtering processes. The results of both objective and subjective tests verify that the proposed algorithm has slightly better quality than the conventional one.*

### I. 서론

최근 다양한 음성통신 서비스에 대한 요구가 증가함에 따라, 서로 다른 음성부호화기를 사용하는 통신망 사이에서는 음성부호화기간의 효율적인 연동이 필요하다. 일반적으로 이러한 음성통신 시스템에선 연산량을 감소

시키고, 음질을 향상시키기 위하여 상호부호화기를 사용한다[1][4]. 상호부호화기는 송신단의 음성부호화기와 수신단의 음성부호화기로 구성된다. 즉, 송신단의 후처리 필터와 수신단의 인지가중 필터가 직렬 연결되기 때문에 서로 상반되는 역할을 하는 두 개의 필터과정을 수행한다. 이러한 상호부호화기의 구조는 불필요한 연산량과 음질 저하를 초래한다. 본 논문은 두 필터의 특성을 이용하여, 두 필터의 역할을 수행하는 하나의 필터를 설계하는 방법을 제안하고 성능을 평가한다.

## II. 상호부호화기

### 2.1 상호부호화기의 구조 및 역할

유/무선 통신망간의 음성 통신을 위해서는 서로 다른 두 음성부호화기가 함께 동작해야 하는 경우가 발생하고, 이러한 경우 두 번의 부호화 과정과 복호화 과정을 거쳐야 한다 (tandem). 그러나, 이러한 tandem방법은 음질저하 및 연산량과 지연시간이 증가하는 단점들이 있다[1][4]. 이러한 이유로 완전한 복호화, 부호화 과정을 거치지 않고, 가능한 경우 비트열에서 비트열로 직접 변환하는 방법이 바람직하다. 이러한 방법을 상호부호화 (transcoding)라고 한다. 상호부호화기의 구조는 송신단의 복호화기와 수신단의 부호화기로 구성되고, LSP 직접변환, 개회로 피치 변환, 고속 적응 코덱북 검색 방

법 등 많은 상호부호화 알고리즘을 사용함으로써 음질 저하와, 연산량을 감소시킬 수 있다[1][4]. 하지만, 일반적인 상호부호화 알고리즘에서는 후처리 필터와 인지가중 필터를 고려하지 않는다. 상호부호화기 내에서 직렬 연결된 구조를 갖는 두 필터는 서로 상반된 주파수 특성을 갖는다. 이러한 상호부호화기의 특징은 불필요한 연산량과 음질저하를 가져 올 수 있다.

### 2.2 인지가중 필터와 후처리 필터

대부분의 음성부호화 표준안은 부호화 과정에서 인지가중 필터를 사용한다. 인지가중 필터는 인간의 청각적 특성을 고려하여 주파수 영역에서의 포먼트 (formant)영역을 약화시키고, valley영역을 강화시킨다. 이러한 과정은 포먼트 영역에 많은 왜곡을 허락하고, 반대로 valley 영역에서의 왜곡을 감소시키는 결과를 가져온다[5]. 일반적인 인지가중 필터를 식 (1)과 같이 나타낼 수 있다.

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)} \quad (1)$$

여기서  $A(z)$ 는 선형예측 필터이고,  $0 < \gamma_1, \gamma_2 < 1$ 는 인지가중 필터의 가중계수들이다. 이 가중계수 ( $\gamma_1, \gamma_2$ )값은 주파수 영역에서의 포먼트 영역을 약화시키고, valley영역을 강화시키는 정도를 조절하고, 각 음성부호화기에 따라 조금씩은 다르지만 거의 비슷한 값을 갖는다 [2][3]. 후처리 필터는 복호화기의 마지막 과정으로, 일반적으로 인지가중 필터에 인한 영향을 보상해 주는 역할을 한다[5][6]. 즉, 인지가중 필터에 의해 약화되었던 포먼트 영역을 강화시키고 valley영역을 약화시키는 역할을 한다. 일반적인 적응 후처리 필터는 다음과 같이 표현된다.

$$H_r(z) = H_f(z)H_i(z) \quad (2)$$

where  $H_f = \frac{\hat{A}(z/\gamma_r)}{\hat{A}(z/\gamma_d)}$  and  $H_i = 1 - \mu z^{-1}$

여기서  $H_i$ 는 tilt보상을 위한 필터로서,  $\mu$ 는 각 서브 프레임 단위로 변하는 계수이다.

### III. 새로운 필터 설계 알고리즘

앞장에서 언급하였듯이 상호부호화기의 인지가중 필터와 후처리 필터는 직렬 연결된 구조를 가지고, 주파수 영역에서 서로 상반되는 역할을 한다. 이러한 필터과정에서 두 필터의 비선형 위상특성으로 인하여 바람직하

지 못한 음성 왜곡이 발생하게 되고, 불필요한 연산량을 가져 온다. 특히, 복호화기의 후처리 필터는 상대적으로 많은 연산량을 필요로 한다[2][3]. 본 절에서는 이러한 문제를 해결하기 위해서 후처리 필터와 인지가중 필터의 역할을 대신 할 수 있는 하나의 새로운 필터를 설계한다. 일반적으로 적응 후처리 필터는 tilt 보상필터를 포함하지만, 객관적 음질 평가와 주관적 음질 평가를 통하여 tilt 보상필터는 음질에 거의 영향을 미치지 않는 것을 확인하였다. 이는 상호부호화기에서는 최종 수신단에 같은 역할을 하는 후처리 필터가 존재하기 때문이다. 따라서, 후처리 필터와 인지가중 필터의 직렬 연결된 필터는 아래와 같은 식으로 나타낼 수 있다.

$$H_R(z) = H_f(z)W(z) = \frac{\hat{A}(z/\gamma_r)}{\hat{A}(z/\gamma_d)} \cdot \frac{A(z/\gamma_1)}{A(z/\gamma_2)} \quad (3)$$

여기서,  $H_R(z)$ 는 후처리 필터와 인지가중 필터의 직렬 연결된 필터를 나타내며, 이후부터는 기준 필터라고 명하기로 한다. 상호부호화기에서는 복호화단의 LSP를 직접 또는 선형 보간하여 부호화단에서 사용한다[1][4]. 따라서, 식 (3)에서  $\hat{A}$ 와  $A$ 는 같고, 각 필터의 극점과 영점의 위치는 같다고 가정할 수 있고, 식 (4)와 같이 다시 쓸 수 있다.

$$H_R(z) = \frac{\prod_{i=1}^p (1 - a_i e^{j\omega_i} z^{-1}) \prod_{i=1}^p (1 - c_i e^{j\omega_i} z^{-1})}{\prod_{i=1}^p (1 - b_i e^{j\omega_i} z^{-1}) \prod_{i=1}^p (1 - d_i e^{j\omega_i} z^{-1})} \quad (4)$$

여기서  $p$ 는 필터의 차수이고,  $a_i, b_i, c_i, d_i$ 는 실수이다. 식 (4)에서 볼 수 있듯이 상호부호화기에서의 후처리 필터와 인지가중 필터의 극점과 영점은 서로 같다고 할 수 있다. 따라서 각도  $\omega_i$ 는 두 필터 모두에서 같다고 할 수 있다. 따라서, 기준 필터와 같은 주파수 크기응답 특성을 가지는 새로운 필터를 설계할 수 있고 식 (5)와 같이 나타낼 수 있다.

$$H_{NEW}(z) = \frac{\prod_{i=1}^p (1 - n_i e^{j\omega_i} z^{-1})}{\prod_{i=1}^p (1 - m_i e^{j\omega_i} z^{-1})} \approx H_f(z)W(z) = H_R(z) \quad (5)$$

여기서  $n_i$ 과  $m_i$ 는 실수이고,  $\omega_i$ 는 식 (4)에서의 같은 각도를 나타내는 값이다. 새로운 필터의 구조는 일반적인 인지가중 필터의 구조와 같도록 구성하였다. 여기서, 제안된 알고리즘의 최종 목표는 기준 필터와 같은 주파

수 크기응답 특성을 갖도록 새로운 필터의 가중계수를 설정하는 것이다. 두 필터의 차수가 다르기 때문에 식 (5)를 만족하는 새로운 필터를 찾는 것은 상당히 어렵다. 하지만, 두 필터의 극점과 영점의 위치가 같다는 조건을 이용하여 극점과 영점의 위치를 고정시키고, 반복적인 알고리즘으로 기준필터와 거의 유사한 주파수 크기응답 특성을 가지는 새로운 필터를 찾을 수 있다. 새로운 필터를 설계하는 과정을 요약하면 아래와 같다.

1. 기준 필터와 새로운 필터의 주파수 크기 응답 특성의 차이를 오차로 정의 한다.
2. 제안된 필터의 가중치 계수를 변화시킨다.
3. 최소의 MSE (Mean Square Error)를 가지는 가중 계수를 찾는다.
4. 선택된 가중 계수를 새로운 필터에 적용한다.
5. 설계된 필터로 상호부호화기의 후처리 필터와 인지가중 필터를 대신한다.

그림 1은 기준 필터와 제안된 알고리즘으로 설계된 필터의 주파수 크기응답 특성을 비교하여 보여준다. 제안된 필터의 주파수 크기응답은 기준 필터와 거의 유사함을 알 수 있다. 그림 2는 제안된 필터 ("proposed")와 상호부호화기 ("conventional")의 후처리 필터와 인지가중 필터를 통과하기 전과 통과한 후의 음성의 주파수 위상응답의 차이를 비교하여 보여준다. 제안된 필터를 사용한 경우가 주파수 위상 왜곡이 더욱 적음을 알 수 있다. 이러한 이유는 비선형 위상특성을 가지는 두 필터를 하나의 필터로 대체함으로써 얻어지는 이득이다. 또한 인지가중 필터의 구조를 가지는 제안된 필터를 사용함으로써 일반적인 상호부호화기에 적용하기가 쉽고, 후처리 필터의 연산량을 감소시킬 수 있다.

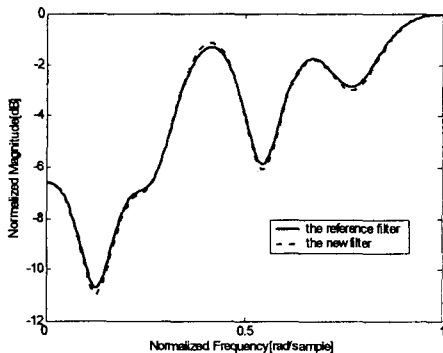


그림 1. 주파수 크기응답 특성 비교

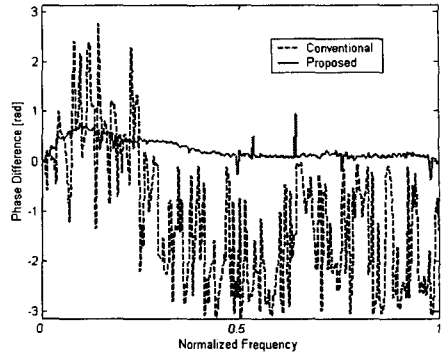


그림 2. 주파수 위상응답의 차이 비교

#### IV. 실험 및 성능평가

본 논문에서는 제안된 알고리즘의 성능을 평가하기 위하여 제안된 필터를 LSP직접변환, 피치 완만화를 이용한 개회로 피치 추정, 고속 적응 코드북 검색 알고리즘을 사용한 AMR과 G.723.1간의 상호부호화기에 적용하였다[1][4]. 객관적 음질평가와 연산량 평가를 위하여 PESQ[7]와 WMOPS[8]를 측정하여 평가하였다.

##### 4.1 연산량 평가

일반적인 상호부호화기와 제안된 알고리즘의 연산량을 tandem방법과 비교하기 위하여 필터과정에 관한 연산량만을 고려하여 WMOPS[8]를 측정하여 표 1에 나타내었다. "Tandem"은 tandem 방법, "Conventional"은 일반적인 상호부호화기, "Proposed"는 제안된 알고리즘의 연산량을 의미한다. "Ratio"은 tandem 방법에 대한 상대적 연산량의 비율을 나타낸다. 표 1에서 보여 주듯이 일반적인 상호부호화기의 경우는 tandem방법에 비하여 필터과정에 관한 연산량에서는 이득이 없음을 알 수 있다. 하지만 제안된 알고리즘의 경우는 두 필터를 하나의 필터로 대체함으로써 약 72%의 연산량이 감소하였다.

##### 4.2 객관적 음질평가

객관적 음질평가를 위하여 8kHz로 표본화된 8초 길이의 NTT Data중 남성, 여성 각각 12명의 음성을 사용하였다. 그림 3은 각 방법에 대한 PESQ[7]를 측정한 결과를 보여준다. "test1"은 상호부호화기에서 인지가중 필터와 후처리 필터 모두를 제거한 경우이고, "test2"는 인

지가중 필터만을 제거한 경우이다. 일반적 상호부화기 (“Conventional”)와 제안된 알고리즘 (“Proposed”)에 비하여 “test1”와 “test2”는 낮은 PESQ값을 보여주는 것을 확인할 수 있다. 이는 상호부화기에서 연산량을 줄이기 위하여 단순히 인지가중 필터와 후처리 필터를 제거하는 것은 많은 음질저하를 초래한다는 것을 의미한다. 실험 결과를 통하여 제안된 알고리즘이 일반적인 상호부화기 보다 다소 향상된 객관적 음질을 보여줌을 확인할 수 있다.

표 1. 연산량 비교 [WMOPS]

	Tandem	Conventional	Proposed
WMOPS	0.691	0.691	0.195
Ratio	100%	100%	28.2%

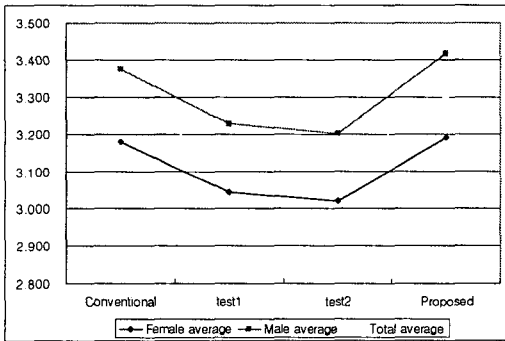


그림 3. 객관적 음질평가 비교 [PESQ]

### 4.3 주관적 음질평가

주관적 음질평가를 위하여 NTT data 남성, 여성 음성 샘플 12개를 사용하여 A-B 선호도 평가를 하였다. 22명의 청자들은 헤드폰을 통하여 두 음성을 듣고, 선호하는 음성 샘플을 선택하였다. 표 2는 주관적 음질평가의 결과를 보여준다. 주관적 음질평가는 35.24%가 제안된 알고리즘을, 29.36%가 일반적 상호부화기 알고리즘을 선호함으로써, 제안된 알고리즘이 기존의 알고리즘보다 다소 높은 선호도를 보였다. 이는 객관적 음질평가의 결과와 일치하는 결과이다.

표 2. 주관적 음질평가 비교

	Conventional	No Preference	Proposed
Female	31.44%	32.58%	35.98%
Male	27.27%	38.26%	34.47%
Total	29.36%	35.42%	35.23%

## V. 결론

본 논문에서는 상호부화기에서의 후처리 필터와 인지가중 필터의 역할을 대신할 수 있는 하나의 새로운 필터를 설계하였다. 새로운 필터는 상호부화기 알고리즘의 특징을 이용하여 인지가중 필터의 구조를 가지도록 설계하였고, AMR과 G.723.1간의 상호부화기에 적용하여 성능 평가를 하였다. 성능평가 결과 제안된 알고리즘은 기존의 상호부화기 알고리즘에 비해 72%의 연산량이 감소되었고, 객관적 음질평가와 주관적 음질평가에서도 다소 향상된 결과를 얻었다. 또한, 제안된 필터는 일반적인 인지가중 필터와 같은 구조를 갖기 때문에 일반적 상호부화기에 적용하기가 용이한 장점을 가진다.

## 참고문헌

- [1] S.W. Yoon, S.K. Jung, Y.C. Park, and D.H. Youn, “An efficient transcoding algorithm for G.723.1 and G.729A speech coders,” in *Proc. Eurospeech 2001*, pp. 2499~2502, Sep 2001.
- [2] ITU-T Rec. G.723.1 “Dual-rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s,” 1996.
- [3] 3GPP TS 26.090 V5.0.0, AMR speech codec; Transcoding functions, Jun 2002.
- [4] K.T. Kim, S.K. Jung, Y.C. Park, Y.S. Choi, D.H. Youn, “An efficient transcoding algorithm for G.723.1 and EVRC speech coders,” in *Proc. IEEE VTS 54th Vehicular Technology Conference (VTC 2001)*, vol.3, pp.1561-1564, Oct.7-10, 2001.
- [5] A.M.Kondoz, *Digital speech coding for low rate communication system*, John Wiley & Sons, 1994.
- [6] W.B. Kelijn, *Speech coding and synthesis*, Elsevier Science B.V., 1995.
- [7] ITU-T Draft Rec P.862 “Perceptual evaluation of speech quality (PESQ), an objective method of end-to-end speech quality assessment of narrowband telephone networks and speech codecs,” May 2000.
- [8] ITU-T Draft Rec P.191 “Software tools for speech and audio coding standardization,” Nov 2000.