

논문 2015-52-7-14

잡음 추정 알고리즘을 이용한 신뢰성 있는 코드벡터 조합의 선정 방법

(A Selection Method of Reliable Codevectors using Noise Estimation Algorithm)

정 승 모*, 김 무 영**

(Seungmo Jung and Moo Young Kim[Ⓢ])

요 약

배경잡음에 강인한 음성인식을 위한 전처리기로써 음성향상 기법이 요구되고 있다. 코드북 기반의 음성향상 기법은 기존 잡음 추정 알고리즘들과 비교하여 nonstationary 배경잡음 환경에 강인하다는 장점이 있다. 하지만 코드북 정보에 의존적이기 때문에 입력신호와 상관성이 떨어지는 코드벡터의 조합을 사용할 경우 성능이 급격히 떨어진다는 단점이 있다. 본 논문에서는 학습된 음성과 잡음 코드벡터를 조합하는 과정에서 입력신호와 상관성이 떨어지는 코드벡터의 조합을 제거함으로써, Log-Spectral Distortion (LSD)과 Perceptual Evaluation of Speech Quality (PESQ) 관점에서 기존 코드북 기반 알고리즘의 성능을 향상시켰다.

Abstract

Speech enhancement has been required as a preprocessor for a noise robust speech recognition system. Codebook-based Speech Enhancement (CBSE) is highly robust in nonstationary noise environments compared with conventional noise estimation algorithms. However, its performance is severely degraded for the codevector combinations that have lower correlation with the input signal since CBSE depends on the trained codebook information. To overcome this problem, only the reliable codevector combinations are selected to be used to remove the codevector combinations that have lower correlation with input signal. The proposed method produces the improved performance compared to the conventional CBSE in terms of Log-Spectral Distortion (LSD) and Perceptual Evaluation of Speech Quality (PESQ)..

Keywords : speech enhancement, codebook-based speech enhancement, noise estimation algorithm, reliable codevectors

I. 서 론

* 학생회원, ** 정회원, 세종대학교 정보통신공학과
(Department of Information and Communication Engineering, Sejong University)

Ⓢ Corresponding Author(E-mail: mooyoung@sejong.ac.kr)

※ 이 논문은 2014년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(NRF-2012R1A1B3002386).

Received ; March 3, 2015 Revised ; May 18, 2015

Accepted ; June 15, 2015

최근 스마트라는 단어가 붙은 제품에는 자연스럽게 사용자 개개인의 홍채, 지문, 음성 인식 등의 인식 기술들이 탑재되어 개인 정보 보호를 위해 사용되고 있다. 이 중에서 음성 인식 시스템의 가장 주된 목표는 주변 잡음이나 장애물 등에 영향을 받지 않고, 사용자의 음성을 정확히 인식하고 그에 따른 요구를 제공해 주

는 것이다. 따라서 주변 잡음에 오염되지 않은 신호가 입력되어야 하고, 만약 음성 신호가 주변 잡음에 오염되었다면 잡음을 제거하고 음성의 질을 향상할 수 있는 전처리기가 요구된다. 현재 입력 신호의 인식률을 향상시키기 위해 다양한 전처리 알고리즘들^[1~2]이 제안되고 있다. 그 중 코드북 기반 음성향상 기법^[3~4]은 모델 기반 음성향상 기법 중 하나로 사전에 음성과 잡음 신호를 코드북으로 학습하여 데이터베이스로 저장한 후, 입력 신호가 들어오면 코드북을 정보로 사용한다. 따라서, 기존 잡음 추정 알고리즘과 비교하여 non-stationary 배경잡음에 강인하다는 장점이 있다. 그러나 현실에 존재하는 모든 잡음 신호를 데이터베이스로 저장하는 것은 실질적으로 저장 용량의 한계가 있기 때문에, 일반적으로 양자화 된 대표적인 잡음 신호들만을 저장하여 사용한다. 하지만 양자화 과정에서 아웃라이어의 경우 일반적으로 입력 신호와 상관성이 많이 떨어질 수 있기 때문에 코드벡터 간 조합과정에서 성능을 저하시키는 요인이 될 수 있다. 따라서 알고리즘의 전반적인 성능을 저하시키는 아웃라이어의 코드벡터를 제거하고 신뢰할 수 있는 조합만을 사용하기 위한 알고리즘이 요구된다.

알고리즘의 잡음 제거를 위한 필터 추정 과정에서 입력 신호와 학습된 코드벡터의 조합에 따른 likelihood를 구하게 된다. 이러한 likelihood는 높을수록 입력신호와 유사한 코드벡터의 조합을 가지기 때문에 신뢰할 수 있는 조합이라고 판단할 수 있다. 반대로 likelihood의 값이 낮을수록 입력신호와 상관성이 떨어지는 코드벡터의 조합이기 때문에 알고리즘의 성능을 저하시키는 요소로 작용할 수 있다. 본 논문에서는 낮은 likelihood를 가지는 코드벡터의 조합을 제거하고 신뢰성 있는 코드벡터의 조합만을 사용하여 성능을 향상시키는 방법을 제안한다^[5]. 또한 신뢰성 있는 코드벡터 조합의 여부를 판단하기 위해 잡음 추정 기법을 통해 추정된 잡음과 음성 코드북을 이용하여 기준이 되는 likelihood를 만들어 사용한다^[6].

II. 본 론

1. 코드북 기반 음성향상 기법

입력 신호 $\mathbf{z}(n)$ 은 서로 독립적인 분포를 가진 음성 신호 $\mathbf{x}(n)$ 와 잡음 신호 $\mathbf{w}(n)$ 가 합쳐진 신호로 가정하

면 다음과 같이 표현될 수 있다.

$$\mathbf{z}(n) = \mathbf{x}(n) + \mathbf{w}(n) \quad (1)$$

입력신호가 주어졌을 때, 파라미터 세트 θ 의 최소 평균 제곱 오차의 최적화를 위한 수식^[4]은 다음과 같다.

$$\hat{\theta} = E\{\theta | \mathbf{z}\} = \int_{\Theta} \theta \frac{p(\mathbf{z} | \theta) p(\theta)}{p(\mathbf{z})} d\theta \quad (2)$$

파라미터 세트 $\theta = [\mu_x, \mu_w, \alpha_x^2, \alpha_w^2]$ 에서 μ_x 와 μ_w 는 각각 음성과 잡음의 선형 예측 계수를, α_x^2 와 α_w^2 는 각각 음성과 잡음의 계인을 나타낸다. $p(\mathbf{z} | \theta)$ 는 영 평균과 음성 및 잡음의 공분산 행렬 $\mathbb{R}_x + \mathbb{R}_w$ 로 이루어진 가우시안 확률밀도함수이다. $\mathbb{R}_x = \alpha_x^2 (A_x^T A_x)^{-1}$ 일 때, A_x 는 첫 열을 $[1, \mu_{x_1}, \dots, \delta_{x_p}, 0, \dots, 0]^T$ 로 가지는 $N \times N$ 의 하부삼각 Toeplitz 행렬이다. 식 (1)에서 가정으로부터 식 (2)의 $p(\theta)$ 은 다음과 같이 나타낼 수 있다.

$$p(\theta) = p(\mu_x, \alpha_x^2) p(\mu_w, \alpha_w^2) \quad (3)$$

간소화를 위해 코드벡터와 계인 간에 독립적이라고 가정하면 $p(\mu_x, \alpha_x^2) = p(\mu_x) p(\alpha_x^2)$ 로 표현할 수 있다. 또한 입력신호와 μ_x , 그리고 μ_w 로부터 다음의 식을 통해 최대 우도 추정으로부터 근사화 된 음성과 잡음 계인 $\alpha_x^{2, ML}$ 와 $\alpha_w^{2, ML}$ 을 얻을 수 있다.

$$A \begin{bmatrix} \alpha_x^2 \\ \alpha_w^2 \end{bmatrix} = B \quad (4)$$

여기서 A 와 B 는 다음과 같다

$$A = \begin{bmatrix} \left\| \frac{(\hat{A}_x(k))^2}{|Z(k)|^4} \right\| & \left\| \frac{\hat{A}_x(k) \hat{A}_w(k)}{|Z(k)|^4} \right\| \\ \left\| \frac{\hat{A}_x(k) \hat{A}_w(k)}{|Z(k)|^4} \right\| & \left\| \frac{(\hat{A}_w(k))^2}{|Z(k)|^4} \right\| \end{bmatrix}$$

$$B = \begin{bmatrix} \left\| \frac{\hat{A}_x(k)}{|Z(k)|^2} \right\| \\ \left\| \frac{\hat{A}_w(k)}{|Z(k)|^2} \right\| \end{bmatrix} \quad (5)$$

여기서 $\|f(k)\| = \int |f(k)| dk$ 이고 $|Z(k)|^2$ 와 \hat{A}_x , \hat{A}_w

는 각각 입력신호 및 음성, 잡음 코드벡터의 파워 스펙트럼을 의미한다.

식 (4)와 식 (5)로부터 구한 계인을 이용하면 식 (2)는 $p(\mathbf{y}|\theta) \approx p(\mathbf{y}|\theta)\delta(\alpha_x^2 - \alpha_x^{2,ML})\delta(\alpha_w^2 - \alpha_w^{2,ML})$ 가 된다. 여기서 $\delta(\cdot)$ 는 디랙-델타 함수이다. 코드벡터와 계인 간의 독립적인 가정 및 근사화를 통해, 식 (2)는 다음과 같이 다시 쓸 수 있다.

$$\begin{aligned} \hat{\theta} &\approx \int_{\theta} \frac{p(\mathbf{z}|\theta)\delta(\alpha_x^2 - \alpha_x^{2,ML})\delta(\alpha_w^2 - \alpha_w^{2,ML})}{p(\mathbf{z})} \\ &\quad \cdot p(\mu_x)p(\mu_w)p(\alpha_x^2)p(\alpha_w^2)d\theta \\ &= \int_{\theta_x} \int_{\theta_w} \theta' \frac{p(\mathbf{z}|\mu_x, \mu_w, \alpha_x^{2,ML}, \alpha_w^{2,ML})}{p(\mathbf{z})} \\ &\quad \cdot p(\mu_x)p(\mu_w)p(\alpha_x^{2,ML})p(\alpha_w^{2,ML})d\theta_x d\theta_w \end{aligned} \quad (6)$$

여기서 분모 $p(\mathbf{z})$ 는 주변화를 통해 정규화를 위한 항이 된다.

$$\begin{aligned} p(\mathbf{z}) &= \int_{\theta_x} \int_{\theta_w} p(\mathbf{z}|\mu_x, \mu_w, \alpha_x^{2,ML}, \alpha_w^{2,ML}) \\ &\quad \cdot p(\mu_x)p(\mu_w)p(\alpha_x^{2,ML})p(\alpha_w^{2,ML})d\theta_x d\theta_w \end{aligned} \quad (7)$$

식 (6)와 식 (7)의 적분은 수치적분을 통하여

$$\begin{aligned} \hat{\theta} &= \frac{1}{N_x N_w} \sum_{m,n=1}^{N_x, N_w} \theta'_{mn} \\ &\quad \cdot \frac{p(\mathbf{z}|\mu_x^m, \mu_w^n, \alpha_{x,mn}^{2,ML}, \alpha_{w,mn}^{2,ML})p(\alpha_{x,mn}^{2,ML})p(\alpha_{w,mn}^{2,ML})}{p(\mathbf{z})} \end{aligned} \quad (8)$$

여기서

$$\begin{aligned} p(\mathbf{z}) &= \frac{1}{N_x N_w} \sum_{m,n=1}^{N_x, N_w} \\ &\quad \cdot p(\mathbf{z}|\mu_x^m, \mu_w^n, \alpha_{x,mn}^{2,ML}, \alpha_{w,mn}^{2,ML})p(\alpha_{x,mn}^{2,ML})p(\alpha_{w,mn}^{2,ML}) \end{aligned} \quad (9)$$

이고, $\theta'_{mn} = [\mu_x^m, \mu_w^n, \alpha_{x,mn}^{2,ML}, \alpha_{w,mn}^{2,ML}]$ 는 m 번째 음성 코드벡터와 n 번째 잡음 코드벡터의 선형 예측 계수, 그리고 최대 우도로 추정된 계인들로 이루어진 파라미터 세트이다. 식 (8)에서 구한 최종 파라미터 $\hat{\theta} = [\hat{\mu}_x, \hat{\mu}_w, \hat{\alpha}_x^2, \hat{\alpha}_w^2]$ 를 이용하면 다음과 같이 필터를 구성할 수 있다.

$$\hat{H}(k) = \frac{\hat{\alpha}_x^2}{|\hat{\mu}_x(\omega)|^2} / \left(\frac{\hat{\alpha}_x^2}{|\hat{\mu}_x(\omega)|^2} + \frac{\hat{\alpha}_w^2}{|\hat{\mu}_w(\omega)|^2} \right) \quad (10)$$

2. 신뢰성 있는 코드벡터 조합의 선정 방법

코드북 기반 음성향상 기법은 음성과 잡음 신호의 정보를 코드북으로 가지고 있기 때문에 non-stationary

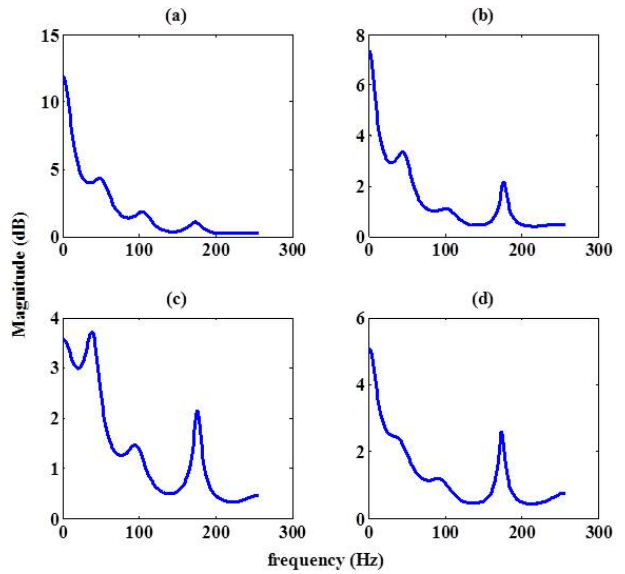


그림 1. 스펙트럼 포락선 비교 (factory1 잡음 0 dB): (a)-(c) 학습된 잡음 코드벡터, (d) 입력된 잡음 신호

Fig. 1. Comparison of spectral envelopes (factory1 noise 0 dB): (a)-(c) trained noise codevectors and (d) input noise signal.

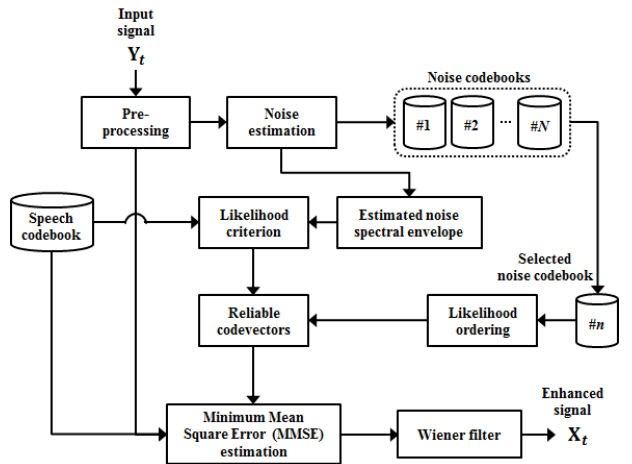


그림 2. 신뢰성 있는 코드벡터 조합 방법의 블록도

Fig. 2. Block diagram of the combination of the reliable codevectors.

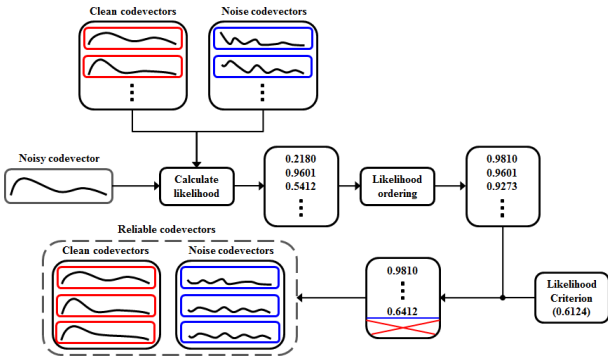


그림 3. 신뢰성 있는 코드벡터 조합의 선정 방법
Fig. 3. The selection method of a reliable codevector combinations.

배경잡음에서도 강인하다는 장점이 있다. 하지만 현실적으로 신호에 대한 모든 정보를 저장하기는 불가능하기 때문에 선형 예측 부호화 된 계수를 양자화 하여 사용한다. 분포된 모든 데이터를 대상으로 균일하게 양자화하기 때문에 빈도수가 적은 학습 데이터 분포에도 양자화가 이루어지지만 상대적으로 적은 수의 codevectors가 생성되며, 이 부분을 아웃라이어라고 한다. 이러한 아웃라이어는 대개 학습 데이터와 상관성이 떨어지는 데이터들로 알고리즘의 성능을 떨어뜨리는 요인으로 작용할 수 있다. 그림 1에서는 실제 입력된 잡음 신호와 학습된 잡음 코드벡터의 스펙트럼 포락선을 비교하였다. 그림 1. (b)와 그림 1. (c)의 경우 그림 1. (d)의 입력신호와 유사한 스펙트럼 포락선을 가지고 있기 때문에 높은 likelihood를 기대할 수 있지만, 그림 1. (a)의 경우 상대적으로 다른 형태의 스펙트럼 포락선을 가지고 있기 때문에, 최종적인 필터링을 구성하는 과정에서 이러한 스펙트럼 포락선들이 알고리즘의 전반적인 성능을 저하시키는 요인이 될 수 있다. 따라서 낮은 likelihood를 갖는 코드벡터 조합을 제거하고, 신뢰성 있는 코드벡터 조합만을 사용하여 성능을 향상시키는 방법을 제안한다. 제안하는 알고리즘의 블록도는 그림 2와 같다. 먼저 입력 신호의 파워 스펙트럼 $|Z_n(k)|^2$ 로부터 잡음 추정 알고리즘을 통해 잡음 신호 $|\tilde{w}_n(k)|^2$ 을 추정한다. 이렇게 추정된 잡음의 파워 스펙트럼을 선형 예측 계수를 통해 $\tilde{\mu}_w$ 를 얻을 수 있다. 추정된 잡음의 선형 예측 계수 $\tilde{\mu}_w$, 음성 코드북, 입력 신호를 이용하여 식 (4)와 (5)를 이용하여 추정된 잡음의 계인 $\beta_w^{2,ML}$ 을 구할 수 있고, 다음의 식을 통해 기준 likelihood $\hat{\mathbf{L}}$

을 얻을 수 있다^[6].

$$\hat{\mathbf{L}} = \frac{1}{N_x} \sum_{m=1}^{N_x} p(\mathbf{z} | \mu_x^m, \tilde{\mu}_w, \alpha_{x,m}^{2,ML}, \beta_w^{2,ML}) \quad (11)$$

그림 3은 제안하는 신뢰성 있는 코드벡터 조합의 선정 방법을 나타낸 그림으로, 훈련된 코드벡터를 이용하여 조합에 따른 likelihood를 다음과 같이 구한다^[5].

$$\mathbf{L}_{(m-1) \times N_w + n} = p(\mathbf{z} | \mu_x^m, \mu_w^n, \alpha_{x,mn}^2, \alpha_{w,mn}^2) \quad (12)$$

식 (12)를 통해 구한 likelihood를 높은 값부터 차례대로 내림차순으로 정렬하면 다음과 같이 내림차순으로 정렬된 likelihood와 해당 likelihood로 구성되는 음성과 잡음의 코드벡터 조합을 얻을 수 있다.

$$\mathbf{L}_{desc} = [\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \dots, \mathbf{L}_{N_x N_w}] \quad (13)$$

가장 위에 정렬된 likelihood는 입력 신호를 가장 잘 나타낼 수 있는 음성과 잡음의 코드벡터의 조합이기 때문에 알고리즘의 성능 향상을 기대할 수 있다. 하지만 likelihood의 값이 떨어질수록 입력 신호와 가장 상관성이 떨어지기 때문에 오히려 성능 저하를 불러올 수 있다. 이렇게 입력 신호와 상관성이 떨어져 알고리즘의 전반적인 성능을 저하시키는 코드벡터의 조합을 배제하고 새로이 신뢰할 수 있는 코드벡터의 조합만을 구성하기 위해 식 (11)에서 구한 기준 likelihood를 이용한다.

$$\mathbf{L}_{rel} = [\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \dots, \mathbf{L}_{\hat{N}_x \hat{N}_w}], \text{ for } \mathbf{L}_{desc} > \hat{\mathbf{L}} \quad (14)$$

여기서 \hat{N}_x , \hat{N}_w 는 기준 likelihood보다 높은 likelihood들의 개수를 의미한다. 최종적으로 기준 likelihood보다 높은 값을 가지는 likelihood를 이루는 코드벡터의 조합은 신뢰할 수 있는 조합이라 판단, 해당 조합을 이루는 파라미터 $\theta_{rel} = [\mu_{x,rel}, \mu_{w,rel}, \alpha_{x,rel}^2, \alpha_{w,rel}^2]$ 만을 이용하여 식 (8), (9)에 적용하면 다음과 같다.

$$\hat{\theta}_{rel} = \frac{1}{\hat{N}_x \hat{N}_w} \sum_{m,n=1}^{\hat{N}_x, \hat{N}_w} \theta'_{mn,rel} \cdot \frac{p(\mathbf{z} | \mu_{x,rel}^m, \mu_{w,rel}^n, \alpha_{x,mn,rel}^{2,ML}, \alpha_{w,mn,rel}^{2,ML}) p(\alpha_{x,mn,rel}^{2,ML}) p(\alpha_{w,mn,rel}^{2,ML})}{p(\mathbf{z}_{rel})} \quad (15)$$

여기서 $p(\mathbf{z}_{rel})$ 은 다음과 같다.

$$p(\mathbf{z}_{rel}) = \frac{1}{\hat{N}_x \hat{N}_w} \sum_{m,n=1}^{\hat{N}_x, \hat{N}_w}$$

$$\cdot p(\mathbf{z} | \mu_{x,rel}^m, \mu_{w,rel}^n, \alpha_{x,mn,rel}^{2,ML}, \alpha_{w,mn,rel}^{2,ML}) p(\alpha_{x,mn,rel}^{2,ML}) p(\alpha_{w,mn,rel}^{2,ML}) \quad (16)$$

III. 실험

본 논문에서는 코드북 학습을 위해 음성 데이터베이스로 TIMIT^[7] training 데이터 4620 문장 (462명의 화자, 각 10 문장)을 사용하였다. 잡음 데이터베이스로는 NoiseX-92^[8] (babble, f16, factory2, hfchannel, white 잡음)를 사용하였다. 해당 데이터베이스들은 양자화를 위해 음성은 10차 선 스펙트럼 주파수 계수^[9]에 10비트를 할당하였고, 잡음은 babble, f16 잡음에 10차 선 스펙트럼 주파수 계수에 4비트를, factory2, hfchannel,

표 1. 기존 알고리즘과 제안한 알고리즘의 LSD 스코어
Table 1. LSD scores of the conventional and proposed algorithms.

Noise	SNR (dB)	기존 알고리즘	제안하는 알고리즘
babble	0	5.26	5.05
	5	3.86	3.66
f16	0	5.10	4.86
	5	3.63	3.38
factory2	0	3.93	3.53
	5	2.78	2.58
hfchannel	0	5.25	4.84
	5	4.43	3.88
white	0	6.83	5.28
	5	5.04	4.23
Average		4.61	4.13

표 2. 기존 알고리즘과 제안한 알고리즘의 PESQ 스코어
Table 2. PESQ scores of the conventional and proposed algorithms.

Noise	SNR (dB)	기존 알고리즘	제안하는 알고리즘
babble	0	2.02	2.06
	5	2.38	2.41
f16	0	2.00	2.17
	5	2.37	2.53
factory2	0	2.05	2.10
	5	2.42	2.45
hfchannel	0	1.71	1.92
	5	1.99	2.20
white	0	2.08	2.10
	5	2.40	2.42
Average		2.14	2.24

white 잡음에 6차 선 스펙트럼 계수에 3비트를 할당하였다.

평가 데이터는 training 데이터에 사용되지 않은 음성 데이터베이스 TIMIT test 데이터 110문장 (11명의 화자, 각 10 문장)을 사용하였고, 학습과 평가에 사용된 모든 데이터들은 8 kHz로 다운 샘플링 되었다. 기준 likelihood를 구하기 위해서는 잡음 추정 알고리즘이 요구되는데, 본 논문에서는 Improved Minima Controlled Recursive Averaging (IMCRA)^[10] 알고리즘을 이용하여 잡음을 추정하였다. 표 1과 2는 기존 코드북 기반 알고리즘^[4]과 제안된 신뢰성 있는 코드벡터만을 사용한 알고리즘을 각각 Log-Spectral Distortion (LSD)과 Perceptual Evaluation of Speech Quality (PESQ) 관점에서 비교하였다. 제안하는 알고리즘이 기존 알고리즘과 비교하여 평균 약 0.48 dB LSD 점수와 약 0.1의 PESQ 성능 향상을 보여주고 있다. 이러한 결과는 코드북이 입력 신호에 대한 사전 정보를 포함하고 있다고 해도 양자화를 통해 입력 신호와 상관성이 떨어지는 코드벡터들을 조합할 경우 오히려 성능을 저하시킬 수 있기 때문에, 해당 조합들을 제거하고 신뢰할 수 있는 코드벡터의 조합만을 사용함으로써 좀 더 향상된 성능을 얻을 수 있다는 것을 보여준다.

IV. 결론

입력 신호의 사전 정보를 가지는 코드북에서 양자화로 인하여 입력 신호와 상관성이 떨어지는 코드벡터들의 조합에 의해 전반적인 성능이 저하되는 문제를 고려하였다. 따라서, 해당 코드벡터의 조합을 배제하고 전반적으로 상관성이 높은 코드벡터만을 조합함으로써 성능을 향상시키는 방법으로 접근하였고, 상관성이 높은 코드벡터만을 선별하기 위한 선정 방법을 제안하였다. 결과적으로 상관성이 높은 코드벡터의 조합만을 사용함으로써 기존 알고리즘과 비교하여 전반적인 성능 향상의 결과를 볼 수 있었다. 그러나 이러한 방법은 장기적으로는 기존 코드북의 근본적인 문제를 해결할 수 없기 때문에 코드북의 저장 용량의 한계, 입력 신호와의 상관성 있는 코드벡터의 학습이 앞으로의 과제이다.

REFERENCES

- [1] P. Loizou, Speech Enhancement: Theory and Practice. CRC Press, 2007.
- [2] S. Jung and M. Y. Kim, "Gain Compensation Method for Codebook-Based Speech Enhancement," Journal of The Institute of Electronics and Information Engineers, Vol. 51, pp. 2051-2056, 2014
- [3] S. Srinivasan, J. Samuelsson, and W. B. Kleijn, "Codebook driven short-term predictor parameter estimation for speech enhancement," IEEE Trans. Speech Audio Process., vol. 14, pp. 163-176, 2006.
- [4] S. Srinivasan, J. Samuelsson, and W. B. Kleijn, "Codebook-Based Bayesian Speech Enhancement for Nonstationary Environments," IEEE Trans. Audio, Speech, Lang. Process., vol. 15, pp. 441-452, 2007.
- [5] I. Hwang, K. Byun, and M. Y. Kim, "Reliable Codevector Selection for Codebook-based Speech Enhancement", KSCSP2011, vol. 28, pp. 267-268, 2011.
- [6] S. Jung and M. Y. Kim, "Reliable Codevectors Selection by eliminating Outlier", KSCSP2014, vol. 31, pp. 163-164, 2014.
- [7] J. Garofolo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, and N. Dahlgren, "DARPA TIMIT acoustic phonetic continuous speech corpus," 1993, CDROM.
- [8] A. Varga, H. J. M. Steeneken, M. Tomlinson, and D. Jones, "The Noisex-92 Study on the Effect of Additive Noise on Automatic Speech Recognition," Technical Report. Malvern, U.K.: DRA Speech Res. Unit, 1992.
- [9] A. D. Subramaniam and B. D. Rao, "PDF optimized parametric vector quantization of speech line spectral frequencies," IEEE Trans. Speech Audio Process., vol. 11, pp. 130-142, 2003.
- [10] I. Cohen, "Noise spectrum estimation in adverse environments: improved minima controlled recursive averaging," IEEE Trans. Speech Audio Process., vol. 11, pp. 466-475, 2003.

저 자 소 개



정 승 모(학생회원)
 2013년 세종대학교 정보통신공학과 학사 졸업
 2015년 세종대학교 정보통신학과 석사 졸업
 <주관심분야 : 잡음 제거, 음성/화자 인식, 음성 신호처리 및 코딩>



김 무 영(정회원)-교신저자
 1993년 연세대학교 전자공학과 학사 졸업
 1995년 연세대학교 전자공학과 석사 졸업
 1995년~2000년 삼성종합기술원 전문연구원
 2001년~2004년 Royal Institute of Technology (KTH, 스웨덴) Dept. Signals, Sensors, Systems, 박사 졸업
 2004년~2005년 Royal Institute of Technology (KTH, 스웨덴) Dept. Signals, Sensors, Systems, PostDoc
 2005년~2006년 Ericsson Research (스웨덴), Senior Research Engineer
 2006년~현재 세종대학교 정보통신공학과, 부교수
 <주관심분야 : 음악정보검색, 음성/오디오 신호처리 및 코딩, 패턴인식, 정보이론.>