

논문 2014-51-9-19

코드북 기반 음성향상 기법을 위한 게인 보상 방법 (Gain Compensation Method for Codebook-Based Speech Enhancement)

정 승 모*, 김 무 영**

(Seungmo Jung and Moo Young Kim[©])

요 약

음성 인식을 위한 전처리기로 주변 잡음을 제거해 주는 음성향상 기법이 강조되고 있다. 다양한 음성향상 기법들 중 코드북 기반 음성향상 기법은 nonstationary 잡음 환경에서도 효율적으로 동작한다. 하지만, 기존 코드북 기반 음성향상 기법에서는 입력 신호와 음성 및 잡음 코드벡터 간에 미스매치가 발생하여 부정확한 게인이 추정되는 문제가 있다. 본 논문에서는 부정확한 게인을 보상하기 위해 long-term 잡음 추정 알고리즘을 사용하여 매 프레임 별로 신호 대 잡음비 기반의 Normalized Weighting Factor (NWF)를 구하고, 이것을 기존 게인에 보상하는 방식을 제안한다. 제안된 코드북 기반 음성향상 기법은 기존 코드북 기반 음성향상 기법에 비해 향상된 성능을 보였다.

Abstract

Speech enhancement techniques that remove surrounding noise are stressed to preprocessor of speech recognition. Among the various speech enhancement techniques, Codebook-based Speech Enhancement (CBSE) operates efficiently in non-stationary noise environments. But, CBSE has some problems that inaccurate gains can be estimated if mismatch occur between input noisy signal and trained speech/noise codevectors. In this paper, the Normalized Weighting Factor (NWF) is calculated by long-term noise estimation algorithm based on Signal-to-Noise Ratio, compensated to the conventional inaccurate gains. The proposed CBSE shows better performance than conventional CBSE.

Keywords : Speech Enhancement, Codebook-based Speech Enhancement, Improved Minima Controlled Recursive Averaging, Normalized Weighting Factor

I. 서 론

최근 스마트폰을 포함한 각종 IT 기기의 발전과 더불어 이를 응용한 소프트웨어에 대해서도 지속적인 연

구가 진행되고 있다. 그 중 음성 인식 시스템의 주된 목표는 주변 소음이나 환경에 따른 영향에 구애받지 않고 사용자의 요구에 따른 정확한 인식 결과를 보여주는 것이다. 따라서 인식을 향상을 위해 음성향상 기법을 전처리기로 사용하는 음성 인식 시스템이 강조되고 추세이다. 기본적으로 음성향상 기법에서는 스펙트럼 차감 및 필터링을^[1] 통해 주변 소음을 제거할 수 있고, 크게 non-model 음성향상 기법과 모델 기반의 음성향상 기법 두 가지로 나눌 수 있다.

Non-model 음성향상 기법 중 Minimum Statistics^[2]나 Minima Controlled Recursive Averaging^[3-4] 등의 long-term 잡음 추정 알고리즘들에서는 음성이 존재하

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

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

※ “본 연구는 미래창조과학부 및 정보통신산업진흥원의 대학 IT연구센터 육성지원사업의 연구결과로 수행되었음” (NIPA-2013-H0301-13-4007).

접수일자: 2014년6월23일, 수정일자: 2014년8월11일

수정완료: 2014년8월26일

지 않거나 묵음 구간일 경우 해당 프레임에 잡음만이 존재한다고 판단, 잡음 스펙트럼을 업데이트 하는 방식이다. long-term 잡음 추정 알고리즘의 장점은 빠른 연산과 stationary 잡음환경에서 좋은 성능을 준다는 장점을 가지고 있으나, nonstationary 잡음환경에서는 낮은 성능을 보여준다는 단점이 있다. 반면에 모델 기반의 음성향상 기법은 사전에 훈련된 모델을 정보로 사용한다는 장점을 가지고 있다. 따라서 기존 모델을 사용하지 않는 음성향상 기법이 가지고 있는 한계였던 nonstationary 잡음환경에서도 좋은 성능을 보여준다. 코드북 기반의 음성향상 기법^[5-6]은 모델 기반 음성향상 기법 중 대표적인 기법으로 음성 및 잡음의 파형을 선형 예측 계수로 바꿔 코드북으로 저장, 코드북으로부터 최대 우도 기반의 추정 기법을 이용하여 계인을 추정한다. 그러나 입력 신호와 음성 및 잡음 코드벡터 간의 미스매치가 커지게 되면 정확한 계인의 추정이 어려워지고, 결과적으로 성능이 저하된다.

본 논문에서는 코드북 기반의 음성향상 기법에서 계인의 부정확한 추정을 줄이기 위해, long-term 잡음 추정 알고리즘으로부터 추정된 음성 및 잡음 신호를 이용하여 신호 대 잡음비 기반의 Normalized Weighting Factor (NWF)를 구하고 이러한 NWF를 기존의 추정된 계인에 보상하는 알고리즘을 제안한다^[7].

II. 본 론

1. 코드북 기반 음성향상 기법

입력 신호 \mathbf{y} 가 주어졌을 때, 파라미터 세트 λ 의 최소 평균 제곱 오차를 최적화하는 수식^[6]은 다음과 같이 쓸 수 있다.

$$\begin{aligned} \hat{\lambda} &= E\{\lambda|\mathbf{y}\} = \int_{\Lambda} \lambda p(\lambda|\mathbf{y})d\lambda \\ &= \int_{\Lambda} \lambda \frac{p(\mathbf{y}|\lambda)p(\lambda)}{p(\mathbf{y})}d\lambda \end{aligned} \quad (1)$$

여기서 파라미터 세트 $\lambda = [\delta_x, \delta_d, g_x^2, g_d^2]$ 에서 δ_x, g_x^2 는 음성의 선형 예측 계수와 계인을, δ_d, g_d^2 는 잡음의 선형 예측 계수와 계인을 나타낸다. $p(\mathbf{y}|\lambda)$ 는 제로 평균과 음성과 잡음의 공분산 행렬 $R_x + R_d$ 을 갖는 가우시안 확률밀도함수이다. $R_x = g_x^2(\Psi_x^T \Psi_x)^{-1}$ 일 때, Ψ_x 는 첫

열을 $\delta_x = [1\delta_{x_1}, \dots, \delta_{x_p}, 0, \dots, 0]^T$ 로 가지는 $N \times N$ 의 하부삼각 Toeplitz 행렬이다. 음성 및 잡음의 분포가 독립적이라는 가정으로부터 식 (1)의 $p(\lambda)$ 은

$$p(\lambda) = p(\delta_x, g_x^2) p(\delta_d, g_d^2) \quad (2)$$

로 나타낼 수 있고, 간소화를 위해 코드벡터와 계인 간에 독립적이라고 가정하면 음성의 코드벡터와 계인은 다음과 같이 $p(\delta_x, g_x^2) = p(\delta_x)p(g_x^2)$ 로 표현할 수 있다. 또한, δ_x 와 δ_d , 그리고 입력 신호가 주어졌을 때, 최대 우도 추정에 의해 음성과 잡음 계인 $g_x^{2,ML}$ 와 $g_d^{2,ML}$ 을 얻을 수 있다^[5-6].

$$U \begin{bmatrix} g_x^2 \\ g_d^2 \end{bmatrix} = V \quad (3)$$

U 와 V 는 각각 다음과 같이 구할 수 있다.

$$\begin{aligned} U &= \begin{bmatrix} \left\| \frac{1}{P_y^2(\omega)|\Psi_x(\omega)|^4} \right\| & \left\| \frac{1}{P_y^2(\omega)|\Psi_x(\omega)|^2|\Psi_d(\omega)|^2} \right\| \\ \left\| \frac{1}{P_y^2(\omega)|\Psi_x(\omega)|_d^2(\omega)^2} \right\| & \left\| \frac{1}{P_y^2(\omega)|\Psi_d(\omega)|^4} \right\| \end{bmatrix} \\ V &= \begin{bmatrix} \left\| \frac{1}{P_y(\omega)|\Psi_x(\omega)|^2} \right\| \\ \left\| \frac{1}{P_y(\omega)|\Psi_d(\omega)|^2} \right\| \end{bmatrix} \end{aligned} \quad (4)$$

여기서 $\|f(\omega)\| = \int |f(\omega)|d\omega$ 이고 $p_y(\omega)$ 는 입력 신호의 파워 스펙트럼이다.

식 (3)과 식 (4)의 최대 우도 추정에 의해 음성 및 잡음의 계인 $g_x^{2,ML}$ 와 $g_d^{2,ML}$ 을 구하면, 식 (1)에서의 $p(\mathbf{y}|\lambda)$ 는 $p(\mathbf{y}|\lambda)\delta(g_x^2 - g_x^{2,ML})\delta(g_d^2 - g_d^{2,ML})$ 로 근사화 된다. 여기서 $\delta(\cdot)$ 는 디랙-델타 함수이다. 위에서 언급한 독립적인 가정과 근사화를 통해 식 (1)은

$$\begin{aligned} \hat{\lambda} &\approx \int_{\Lambda} \lambda \frac{p(\mathbf{y}|\lambda)\delta(g_x^2 - g_x^{2,ML})\delta(g_d^2 - g_d^{2,ML})}{p(\mathbf{y})} \\ &\quad \cdot p(\delta_x)p(\delta_d)p(g_x^2)p(g_d^2)d\lambda \\ &= \int_{\Delta_x} \int_{\Delta_d} \lambda \frac{p(\mathbf{y}|\delta_x, \delta_d, g_x^{2,ML}, g_d^{2,ML})}{p(\mathbf{y})} \\ &\quad \cdot p(\delta_x)p(\delta_d)p(g_x^{2,ML})p(g_d^{2,ML})d\delta_x d\delta_d \end{aligned} \quad (5)$$

으로 다시 쓸 수 있고, 분모 $p(\mathbf{y})$ 는 다음과 같이 주변화를 통해 정규화를 위한 항이 된다.

$$p(\mathbf{y}) = \int_{\Delta_x} \int_{\Delta_d} p(\mathbf{y}|\delta_x, \delta_d, g_x^{2,ML}, g_d^{2,ML}) \cdot p(\delta_x) p(\delta_d) p(g_x^{2,ML}) p(g_d^{2,ML}) d\delta_x d\delta_d \quad (6)$$

식 (5)와 식 (6)의 적분은 수치적분을 통하여

$$\hat{\lambda} = \frac{1}{N_x N_d} \sum_{m,n=1}^{N_x N_d} \lambda'_{mn} \cdot \frac{p(\mathbf{y}|\delta_x^m, \delta_d^n, g_x^{2,ML}, g_d^{2,ML}) p(g_x^{2,ML}) p(g_d^{2,ML})}{p(\mathbf{y})}$$

$$p(\mathbf{y}) = \frac{1}{N_x N_d} \sum_{m,n=1}^{N_x N_d} p(\mathbf{y}|\delta_x^m, \delta_d^n, g_x^{2,ML}, g_d^{2,ML}) p(g_x^{2,ML}) p(g_d^{2,ML}) \quad (7)$$

로 나타낼 수 있고, 여기서 $\lambda'_{mn} = [\delta_x^m, \delta_d^n, g_x^{2,ML}, g_d^{2,ML}]$ 는 m 번째 음성 코드벡터와 n 번째 잡음 코드벡터의 선형 예측 계수, 그리고 이들의 조합에 따른 최대 우도로 추정된 계인으로 이루어진 파라미터 세트이다. 위 식에서 구한 우도와 각 음성과 잡음의 코드벡터가 결합하여 최종적인 최종 평균 제곱 오차 기반의 가중치 합된 파라미터 $\hat{\lambda} = [\hat{\delta}_x, \hat{\delta}_d, \hat{g}_x^2, \hat{g}_d^2]$ 를 구할 수 있고, 이 파라미터를 이용하여 다음과 같이 필터를 구성 할 수 있다^[5~6].

$$H(\omega) = \frac{\hat{g}_x^2}{|\hat{\delta}_x(\omega)|^2} \left/ \left(\frac{\hat{g}_x^2}{|\hat{\delta}_x(\omega)|^2} + \frac{\hat{g}_d^2}{|\hat{\delta}_d(\omega)|^2} \right) \right. \quad (8)$$

2. NWF를 이용한 계인 보상 방법

코드북 기반 음성향상 기법에서 음성과 잡음의 정확한 계인 추정 은 성능에 매우 큰 영향을 준다. 훈련된 음성 및 잡음의 코드벡터와 입력 신호의 코드벡터 간의 스펙트럼이 유사할수록 더욱 정확한 계인을 추정할 수 있다. 그러나 데이터베이스 관점에서의 저장 한계로 인하여 훈련된 코드북은 음성의 대표적인 코드벡터만을 저장하고 있다. 그림 1에서는 훈련된 음성의 코드벡터와 입력 신호 간 미스매치를 최적의 음성 코드벡터와의 스펙트럼 포락선을 통해 비교하였고, 실제 훈련된 잡음의 코드벡터에서도 최적의 잡음 코드벡터와 비교하여

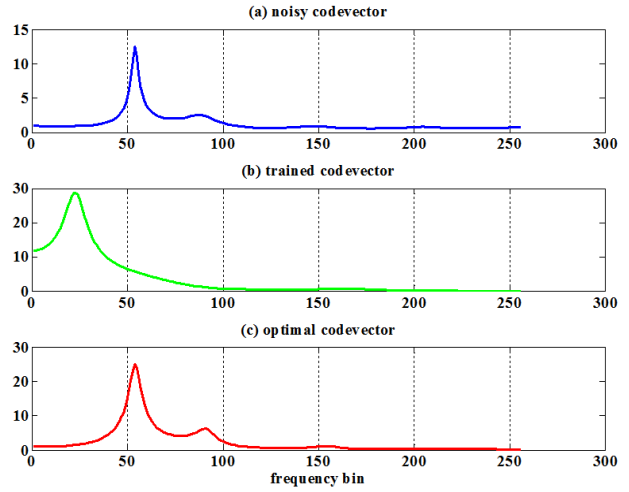


그림 1. 코드벡터 스펙트럼 포락선 비교 (a) 입력 신호, (b) 훈련된 음성, (c) 최적의 음성
Fig. 1. Comparison of spectral envelopes for (a) input noisy signal, (b) trained speech codevector, and (c) optimal speech codevector

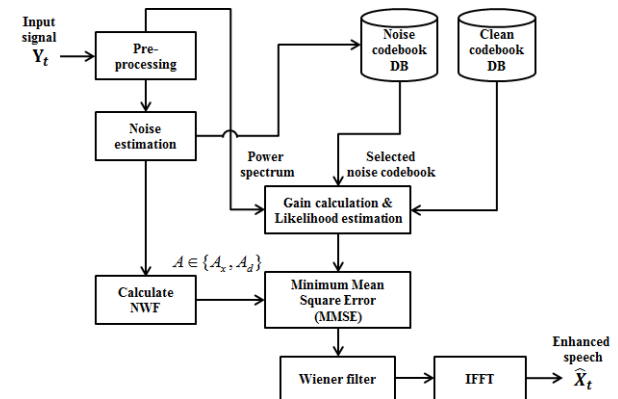


그림 2. 계인 보상된 코드북 기반 음성향상 기법의 블록도
Fig. 2. Block-diagram of gain compensated CBSE.

입력신호와와의 큰 미스매치를 가진다.

본 논문에서는 입력 신호와 훈련된 음성 및 잡음 코드벡터 간의 미스매치로부터 추정되는 부정확한 계인을 보상하는 신호 대 잡음비 기반의 NWF를 제안하고 있다. 그림 2는 본 논문에서 제안하는 알고리즘의 블록도로 식 (1)에 제안하는 NWF를 적용하면 다음과 같이 다시 쓸 수 있다.

$$\hat{\lambda}_{NWF} = \int_{\Lambda(A)} \lambda(A) p(\lambda|\mathbf{y}) d\lambda(A)$$

$$= \int_{\Lambda(A)} \lambda(A) \frac{p(\mathbf{y}|\lambda) p(\lambda)}{p(\mathbf{y})} d\lambda(A) \quad (9)$$

여기서 $\lambda(A)$ 는 기존 파라미터 세트 λ 에 NWF를 적용한 것으로, 먼저 입력 신호로부터 잡음 추정 알고리즘을 이용하여 잡음 신호를 추정한다. 본 논문에서는 Improved Minima Controlled Recursive Averaging (IMCRA) 알고리즘^[4]을 사용하여 입력 신호로부터 잡음 신호 $|\hat{D}_n(k)|^2$ 을 추정하였고, 스펙트럼 차감^[8]을 통해 음성 신호 $|\hat{X}_n(k)|^2$ 을 구하였다.

$$|\hat{X}_n(k)|^2 = \begin{cases} |Y_n(k)|^2 - |\hat{D}_n(k)|^2 & \text{if } |Y_n(k)|^2 - |\hat{D}_n(k)|^2 > \max(|\hat{D}_n(k)|^2) \\ \beta |\hat{D}_n(k)|^2 & \text{else} \end{cases} \quad (10)$$

이렇게 매 프레임 별로 구한 음성 신호 $|\hat{X}_n(k)|^2$ 과 잡음 신호 $|\hat{D}_n(k)|^2$ 을 이용하여 NWF A 를 다음과 같이 구할 수 있다.

$$A = \begin{cases} A_x : \sum_{k=1}^K \frac{|\hat{X}_n(k)|^2}{|\hat{X}_n(k)|^2 + |\hat{D}_n(k)|^2} \\ A_d : \sum_{k=1}^K \frac{|\hat{D}_n(k)|^2}{|\hat{X}_n(k)|^2 + |\hat{D}_n(k)|^2} \end{cases} \quad (11)$$

여기서 K 는 주파수 빈의 수를, A_x 과 A_d 은 각각 음성 및 잡음의 NWF를 나타낸다. 즉, 매 프레임마다 각각 음성과 잡음의 에너지를 이용하여 신호 대 잡음비로 구하고, 정규화를 통해 얻은 값을 기존 개인에 가중치로 보상을 한다. NWF 파라미터를 적용한 식 (9)를 다시 유도하면 다음과 같다.

$$\hat{\lambda}_{NWF} = \frac{1}{N_x N_d} \sum_{m,n=1}^{N_x N_d} \lambda'_{mn}(A) \cdot \frac{p(\mathbf{y}|\lambda'_{mn}(A))p(g_{x,mn}^{2,ML}(A))p(g_{d,mn}^{2,ML}(A))}{p(\mathbf{y})}$$

$$p(\mathbf{y}) = \frac{1}{N_x N_d} \sum_{m,n=1}^{N_x N_d} \cdot p(\mathbf{y}|\lambda'_{mn}(A))p(g_{x,mn}^{2,ML}(A))p(g_{d,mn}^{2,ML}(A)) \quad (12)$$

여기서 $\lambda'_{mn}(A) = [\delta_x^n(A), \delta_d^n(A), g_{x,mn}^{2,ML}(A), g_{d,mn}^{2,ML}(A)]$ 는 기존 λ'_{mn} 에 NWF를 적용한 파라미터 세트이다. 식 (7)에서와 마찬가지로 식 (12)에서 구한 최소 평균 제곱 오차의 가중치 합을 이용하면 NWF가 적용된 필터를 구할 수 있다.

그림 3에서는 각각 최적의 개인, 훈련된 음성 및 잡음

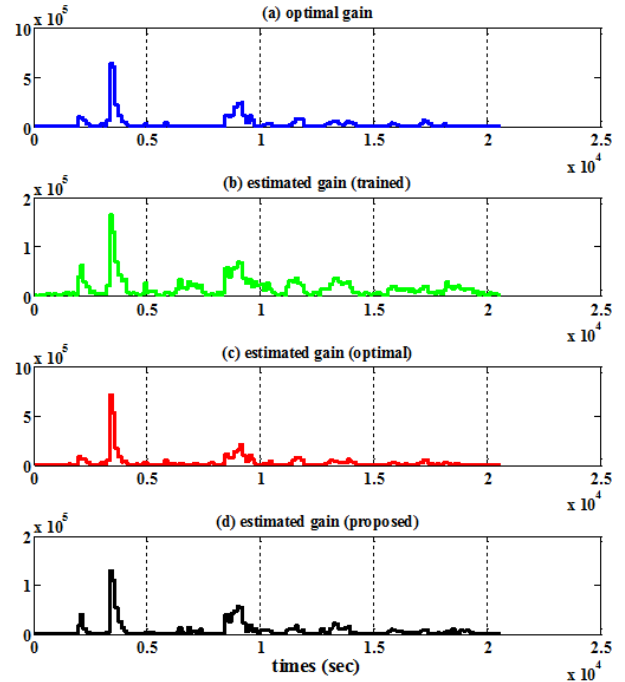


그림 3. 음성의 최적 및 추정 개인의 비교 (5dB white 잡음) (a) 최적의 개인, (b) 훈련된 음성 코드벡터, (c) 최적의 음성 코드벡터, (d) NWF로 보상된 개인

Fig. 3. Comparison of optimal/estimated speech gain (white noise at 5dB) for (a) optimal gain, (b) estimated gain using trained speech codevector, (c) estimated gain using optimal speech codevector, and (d) compensated speech gain using NWF.

음 코드벡터로부터 추정된 개인, 최적의 음성 및 잡음 코드벡터로부터 추정된 개인, NWF를 적용하여 보상된 개인을 보여주고 있다. 결과적으로 최적의 코드벡터를 사용하여 구한 개인보다는 정확하지는 않지만 기존의 부정확한 개인보다는 훨씬 더 최적의 개인에 가까워진다는 것을 볼 수 있다.

III. 실험

코드북을 훈련하는 과정에서 기존 선형 예측 계수로 양자화를 할 경우 계수간의 상관관계가 낮기 때문에 성능이 떨어지는 코드북이 모델링 되는 한계가 있다. 따라서 본 논문의 실험에서는 기존 선형 예측 계수를 선 스펙트럼 주파수^[9]로 변환하여 양자화^[10-11]하였다. 음성 코드북 훈련을 위해서 8 kHz로 다운 샘플링 된 TIMIT 데이터베이스^[12] 트레이닝 파일(462명 화자, 화자 당 10

표 1. 기존 알고리즘과 제안한 알고리즘의 LSD 스코어

Table 1. LSD scores of conventional IMCRA algorithm and proposed CBSE

Noise	SNR (dB)	IMCRA	기존 CBSE	제안하는 CBSE
Babble	0	5.73	5.31	4.75
	5	4.07	3.86	3.65
fl6	0	5.47	5.10	4.44
	5	3.90	3.63	3.41
pink	0	6.02	5.64	4.63
	5	4.30	3.99	3.65
white	0	8.23	6.83	5.30
	5	6.12	5.04	4.35
Average		5.48	4.92	4.27

문장)을 사용하였고, 10차 선 스펙트럼 주파수 계수에 양자화 비트는 10비트를 할당하여 양자화 하였다. 잡음 코드북은 8 kHz로 다운 샘플링 된 NoiseX-92 데이터베이스^[13] (Babble, fl6, pink, white 잡음) 파일을 사용하였다. Babble, fl6 잡음에 대해서는 10차 선 스펙트럼 주파수 계수에 양자화 비트는 4비트를, pink와 white 잡음에 대해서는 6차 선 스펙트럼 주파수 계수에 3비트로 양자화 하였다.

평가 데이터는 트레이닝 데이터와 동일하게 8 kHz로 다운 샘플링 된 TIMIT 데이터베이스 테스트 파일 (11 명의 화자, 화자 당 10문장)을 사용하였다. 잡음 데이터는 Babble, fl6, pink, white 잡음을 사용하였고, 0-5 dB의 신호 대 잡음비로 음성데이터와 혼합되었다. 평가를 위해 기존 IMCRA 알고리즘과 코드북 기반 음성향상 기법, 그리고 제안하는 NWF로 계인을 보상한 코드북 기반 음성향상 기법이 사용되었고, 각각 Log-Spectral Distortion (LSD)과 스펙트로그램 관점에서 비교하였다.

표 1을 보면 IMCRA 알고리즘과 비교하여 기존 코드북 기반 음성향상 기법이 평균 0.56 dB 향상된 LSD 스코어를 보여주고 있다. 앞서 언급한 것과 같이 코드북을 사전 정보로 사용하는 모델 기반의 알고리즘이 IMCRA 알고리즘과 비교하여 nonstationary 잡음 환경에 더 강인하다는 것을 보여주는 결과이다. 제안한 NWF를 적용한 코드북 기반 음성향상 기법은 IMCRA 알고리즘과 비교하여 평균 1.21 dB, 기존 코드북 기반 음성향상 알고리즘과 비교하여 평균 0.65 dB의 향상된 LSD 스코어를 보여준다. 이러한 결과는 기존 음성향상 기법에서 입력 신호와 음성 및 잡음 코드벡터 간 mismatch에 따른 부정확한 계인을 NWF로 보상함으로써,

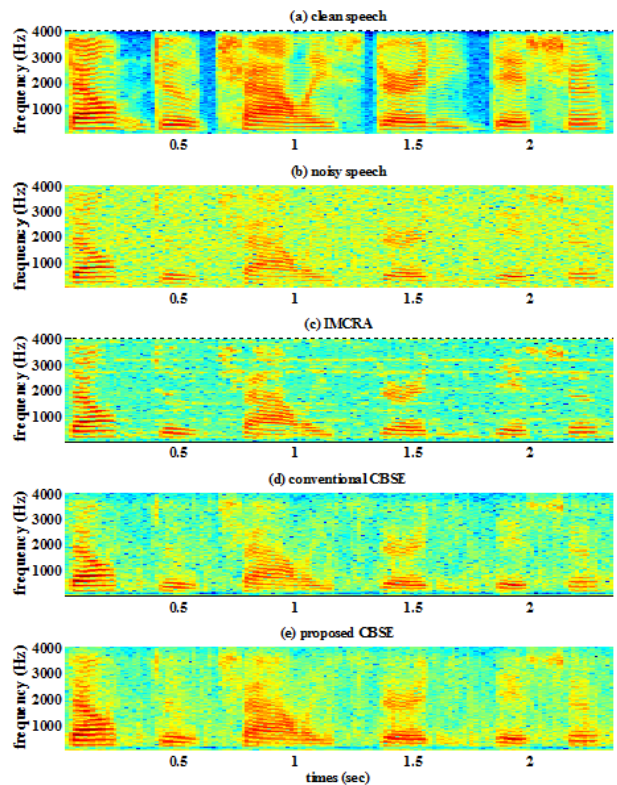


그림 4. 스펙트로그램 결과 (5dB white 잡음) (a) 깨끗한 음성, (b) 입력 신호, (c) IMCRA (d) 기존 CBSE, (e) 제안한 CBSE

Fig. 4. Result of spectrogram (white noise at 5dB) (a) clean speech, (b) noisy signal, (c) IMCRA, (d) conventional CBSE, and (e) proposed CBSE

기존 코드북 기반 음성향상 기법에서 추정된 계인보다 좀 더 최적화 된 계인에 가까운 계인이 추정되었다는 것으로 볼 수 있다.

IV. 결 론

입력 신호와 훈련된 음성 및 잡음 코드벡터 간에 mismatch가 일어날 경우 최적의 코드벡터를 사용했을 때와 비교하여 부정확한 계인의 추정이 구해진다. 이러한 계인의 오차를 IMCRA로 구한 NWF를 적용하여 보상해줌으로써 IMCRA 알고리즘과 비교하여 평균 1.21 dB, 기존 코드북 기반 음성향상 기법과 비교하여 평균 0.65 dB LSD 스코어 향상을 보였다. 그러나 mismatch에 의해 추정된 계인 자체는 여전히 큰 오차를 가지고 있기 때문에 정확한 계인의 추정이 앞으로의 과제이다.

REFERENCES

- [1] R. M. Udea, N. D. Vizireanu, and S. Ciochina, "An improved spectral subtraction method for speech enhancement using a perceptual weighting filter," Elsevier Digital Signal Process., vol. 18, pp. 581 - 587, 2008.
- [2] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," IEEE Trans. Speech Audio Process., vol. 9, pp. 504-512, 2001.
- [3] I. Cohen, "Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement," IEEE Signal Process. Lett., vol. 9, pp. 12-15, 2002.
- [4] I. Cohen, "Noise spectrum estimation in adverse environments: improved minima controlled recursive averaging," IEEE Trans. Speech Audio Process., vol. 11, pp. 466-475, 2003
- [5] 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.
- [6] 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.
- [7] S. Jung and M. Y. Kim, "Codebook-based Speech Enhancement Using an SNR Weighting Factor," The Institute of Electronics Engineers of Korea, pp. 495-496, 2013.
- [8] P. Loizou, Speech Enhancement: Theory and Practice. CRC Press, 2007.
- [9] W. B. Kleijn, T. Backstrom, and P. Alku, "On line spectral frequencies," IEEE Signal Process. Lett., vol. 10, pp. 75-77, 2003
- [10] F. Merazka, "VQ Codebook Design Using Genetic Algorithms for Speech Line Spectral Frequencies," Springer Computational Intelligence and Intelligent Systems, pp. 557-566, 2012.
- [11] 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
- [12] J. Garofolo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, and N. Dahlgren, "DARPA TIMIT acoustic phonetic continuous speech corpus," 1993, CDROM.
- [13] 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.

저자 소개



정 승 모(학생회원)
2013년 세종대학교 정보통신
공학과 학사 졸업
2013년~현재 세종대학교 정보통신학과 석사과정
<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.>