

스펙트럼 변이 기반의 향상된 음성 존재 불확실성 추적 기법을 이용한 Global Soft Decision

Global Soft Decision Based on Improved Speech Presence Uncertainty Tracking Method Incorporating Spectral Gradient

김종웅, 장준혁[†]

(Jong-Woong Kim and Joon-Hyuk Chang[†])

한양대학교 융합전자공학부

(접수일자: 2012년 12월 11일; 수정일자: 2013년 1월 7일; 채택일자: 2013년 1월 28일)

초 록: 본 논문에서는 기존의 global soft decision 기법에서 음성 부재 확률을 구할 때의 음성 부재와 존재에 대한 *a priori* 확률값의 비 (q)에 스펙트럼 변이 기법을 적용한 음성 향상 기법을 제안한다. 기존의 global soft decision 방법은 음성 부재 확률을 구하기 위해 가정한 가설에 따라 고정된 q 값을 사용하였지만, 본 논문에서 제안한 알고리즘은 기존의 고정된 값에 직전 2 프레임에서의 음성 존재 여부와 스펙트럼 변이 값의 상태 조건에 따라 적응적으로 q 값이 가변되도록 하여 음성 부재 확률을 향상시키는 기법이다. 제안된 방법의 성능 평가를 위해 ITU-T P.862 PESQ (Perceptual Evaluation of Speech Quality)를 이용하여 평가하였고, 그 결과 제안된 스펙트럼 변이 기법을 적용한 방법이 기존의 global soft decision 방법보다 향상된 결과를 보여주었다.

핵심용어: Global soft decision, 스펙트럼 변이, q 값

ABSTRACT: In this paper, we propose a novel speech enhancement method to improve the performance of the conventional global soft decision which is based on the spectral gradient method applied to the ratio of *a priori* speech absence and presence probability value (q). Conventional global soft decision scheme used a fixed value of q in accordance with the hypothesis assumed, but the proposed algorithm is a technique for improving the speech absence probability which is applied adaptively variable value of q according to the speech presence or absence in the previous two frames and the conditions of the spectral gradient value. Experimental results show that the proposed improved global soft decision method based on the spectral gradient method yields better results compared to the conventional global soft decision technique based on the performance criteria of the ITU-T P. 862 PESQ (Perceptual Evaluation of Speech Quality).

Keywords: Global soft decision, Spectral gradient, Value of q

PACS numbers: 43.72. Dv

1. 서 론

최근 우리가 살아가는 실생활에서도 실제적인 음성 신호처리 시스템이 필요한 환경이 늘어나면서 음성 향상 기술에 대한 연구가 주목 받고 있다. 실제로 음성 신호처리 기술에서 음성 향상을 위한 다양한 방법들이 시도되었다. 이러한 방법들 중에는 스펙트

럼 차감법,^[1] Wiener filtering,^[2] soft decision 추정,^[3] 최소 평균 자승 오차(MMSE, Minimum Mean Square Error)^[4] 등이 주로 사용되고 있다. 위의 방법들은 구현상의 편리함과 다양한 배경 잡음에 적용 가능한 이점을 가지고 있으며, soft decision에 근거한 음성 향상 방법이 뛰어난 성능을 가진다는 것이 알려져 있다. 특히 global soft decision 방법에서는 기존의 지역 음성 부재 확률(LSAP, Local Speech Absence Probability)과 매 프레임에서 하나의 값으로 표현되어지는 전역

[†]Corresponding author: Joon-Hyuk Chang (jchang@hanyang.ac.kr)
Department of Electronic Engineering, Hanyang University, Seoul, 133-791, Republic of Korea.
(Tel: 82-2-2220-0355, Fax: 82-2-2291-0357)

음성 부재 확률(GSAP, Global Speech Absence Probability)이 결합되어 통계적으로 견실한 음성 부재 확률을 새롭게 도출하였다.^[5]

하지만 음성 부재 확률을 구하기 위해 기존의 통계적 가정을 바탕으로 고정된 음성 부재와 존재에 대한 *a priori* 확률값의 비(q)를 적용한 것은,^[3-5] 다양한 음성 환경 변화에서 한계를 나타낸다. 특히 비정상적인 잡음 환경을 고려하였을 때, 잡음이 음성 신호를 간섭하면 음성의 꼬리와 같은 약한 음성 신호를 추정하기 힘들다. 이를 극복하기 위해 Malah가 제안한 음성 존재 부정확성 추적 방법은 기존의 soft decision에서 음성 부재 확률을 구할 때 사용되는 q 값이 고정된 값이었던 것과는 다르게 프레임, 채널마다 입력 신호의 *a posteriori* SNR을 특정 임계값과 비교하여 음성인지 아닌지를 판별하여 q 값을 다르게 적용하여 음성 부재 확률의 신뢰도를 높이는 방법으로 음성을 향상시켰다.^[6] 또한 이 Malah의 음성 존재 부정확성 추적 방법에서 입력 신호가 음성인지 아닌지 여부를 판별할 때 사용된 *a posteriori* SNR을 대신하여, 잡음이 섞인 신호의 국부 에너지와 주어진 윈도우에서의 최소값 사이의 비를 특정 임계값과 비교하여 음성 유무를 판별 후 q 값을 다르게 적용하여 음성을 향상시키는 방법도 있었다.^[7]

본 논문에서는 기존의 global soft decision에서 음성 부재 확률(SAP, Speech Absence Probability)을 구할 때 사용되는 고정된 q 값이나, 단순히 특정 파라미터 값을 임계값과 비교하여 몇 가지의 다른 q 값을 적용했던 방법과는 달리, 직전 2 프레임에서의 음성 존재 여부와 스펙트럼 변이 값을 기반으로 한 세 가지의 상태 조건에 따른 시그모이드 형태(sigmoid type) 함수를 사용하여 모든 프레임마다 적응적으로 변화하는 각기 다른 q 값을 이끌어냄으로써 보다 견실한 음성 부재 확률을 추정하는 기법을 제안한다. 현재 프레임은 인접한 프레임들에 큰 영향을 받기 때문에 프레임 간 상관관계와 연산량을 고려하여 직전 2 프레임에서의 음성 존재 여부를 활용하였다.^[9] 그리고 스펙트럼 변이는 이전 프레임 동안 예측된 평균 long-term 파워 스펙트럼과 현재 파워 스펙트럼의 차이다.^[8] 즉, 직전 2 프레임이 모두 음성 존재일 경우와 그 이외의 경우를 구분하여,^[9] 각각 프레임 간 스펙

트럼 변이에 따라 증가하는 경우, 유지되는 경우, 감소하는 경우의 세 가지 경우에 대해 q 값을 적응적으로 변화하게 하여 보다 견실한 음성 부재 확률을 도출하는 새로운 알고리즘을 도입하였다.

제안된 음성 향상 기법은 PESQ(Perceptual Evaluation of Speech Quality)와^[10] C_{ov} (composite measure) 방법을^[11] 통해 평가하였고, 기존의 global soft decision 방법보다 향상된 결과를 나타내었다.

II. Global Soft Decision 개요

먼저 원래의 음성 신호 $x(t)$ 에 잡음 신호 $n(t)$ 가 더해져서 오염된 음성 신호 $y(t)$ 를 만들었다고 가정한다. 음성 향상 기법에서 사용되고 있는 기본 가설 $H_0(k, n), H_1(k, n)$ 이 각각 음성의 부재와 존재를 나타낸다고 하면 다음과 같이 표현된다.

$$H_0(k, n): Y(k, n) = N(k, n) \quad (1)$$

$$H_1(k, n): Y(k, n) = X(k, n) + N(k, n) \quad (2)$$

여기서 $Y(k, n), X(k, n)$ 그리고 $N(k, n)$ 은 각각 입력 신호, 원래 음성 신호, 그리고 잡음 신호의 이산 푸리에 변환(DFT, Discrete Fourier Transform) 계수를 나타내고, n 번째 프레임에서의 $k(=0, 1, \dots, K-1)$ 번째 주파수 성분이 된다.

음성 신호와 잡음의 스펙트럼이 복소가우시안 분포를 따른다고 가정을 하면, 가설 $H_0(k, n), H_1(k, n)$ 을 조건으로 한 확률밀도함수는 다음과 같이 주어진다.

$$P(Y(k, n)|H_0) = \frac{1}{\pi\lambda_d(k, n)} \exp\left\{-\frac{|Y(k, n)|^2}{\lambda_d(k, n)}\right\} \quad (3)$$

$$P(Y(k, n)|H_1) = \frac{1}{\pi(\lambda_x(k, n) + \lambda_d(k, n))} \cdot \exp\left\{-\frac{|Y(k, n)|^2}{(\lambda_x(k, n) + \lambda_d(k, n))}\right\} \quad (4)$$

위에서 $\lambda_x(k, n), \lambda_d(k, n)$ 는 각각 음성과 잡음의 분산을 나타낸다. 음성의 존재와 부재에 관한 가설을 바

탕으로 주파수 채널별 지역 음성 부재 확률은 다음과 같이 구해질 수 있다.

$$\begin{aligned} P(H_0|Y(k,n)) &= \frac{P(Y(k,n)|H_0)P(H_0)}{P(Y(k,n))} \\ &= \frac{1}{1 + \frac{P(H_1)}{P(H_0)}\Lambda(Y(k,n))}, \end{aligned} \quad (5)$$

또한 한 프레임에서 음성 부재 확률은 현재 프레임의 관찰 결과를 기반으로 다음과 같이 구할 수 있다.

$$\begin{aligned} P(H_0|Y(n)) &= \frac{P(Y(n)|H_0)P(H_0)}{P(Y(n))} \\ &= \frac{P(Y(n)|H_0)P(H_0)}{P(Y(n)|H_0)P(H_0) + P(Y(n)|H_1)P(H_1)}, \end{aligned} \quad (6)$$

여기서 각 주파수 대역이 통계적으로 서로 독립이라고 가정하면 한 프레임에서의 음성 부재 확률을 다음과 같이 표현할 수 있다.

$$P(H_0|Y(n)) = \frac{1}{1 + \frac{P(H_1)}{P(H_0)} \prod_{k=0}^{K-1} \Lambda(Y(k,n))}, \quad (7)$$

여기서 $P(H_1)/P(H_0)$ 은 음성 부재와 존재에 대한 *a priori* 확률값의 비로서 q 로 나타낼 수 있고,^[5] $\Lambda(Y(k,n))$ 는 k 번째 주파수 채널에서의 우도비 (likelihood ratio)로서 다음과 같이 나타낼 수 있다.

$$\begin{aligned} \Lambda(Y(k,n)) &= \frac{P(Y(k,n)|H_1)}{P(Y(k,n)|H_0)} \\ &= \frac{1}{1 + \xi(k,n)} \exp\left\{ \frac{\gamma(k,n)\xi(k,n)}{1 + \xi(k,n)} \right\}. \end{aligned} \quad (8)$$

여기서 $\xi(k,n) = \frac{\lambda_x(k,n)}{\lambda_d(k,n)}$, $\gamma(k,n) = \frac{|Y(k,n)|^2}{\lambda_d(k,n)}$ 이 되고, $\xi(k,n)$, $\gamma(k,n)$ 는 각각 *a priori* SNR과 *a posteriori* SNR을 나타낸다. 기존의 global soft decision에서 이 q 값은 보통 고정된 값으로 사용되었다. 본 논문에서는 이 q 값을 직전 2 프레임에서의 음성 존재 여부와 스펙트럼 변이 값의 상태 조건에 따라 적응적으로

변화하게 하여 음성 부재 확률을 향상시키는 기법을 제안한다.

또한 음성의 존재와 부재를 고려한 음성과 잡음의 전력 스펙트럼의 평균 기대값은 다음과 같이 주어질 수 있다.^[5]

$$\begin{aligned} E[|N(k,n)|^2|Y(k,n)] &= E[|N(k,n)|^2|Y(k,n), H_0]P(H_0|Y(k,n)) \\ &\quad + E[|N(k,n)|^2|Y(k,n), H_1]P(H_1|Y(k,n)), \end{aligned} \quad (9)$$

$$\begin{aligned} E[|X(k,n)|^2|Y(k,n)] &= E[|X(k,n)|^2|Y(k,n), H_0]P(H_0|Y(k,n)) \\ &\quad + E[|X(k,n)|^2|Y(k,n), H_1]P(H_1|Y(k,n)). \end{aligned} \quad (10)$$

III. 스펙트럼 변이를 이용한 향상된 Global Soft Decision

기존의 global soft decision 방법에서는 고정된 파라미터 값 q 를 사용하였기 때문에 비정상적인 잡음 환경에서 정확한 음성 부재 확률을 추정하지 못하였다. 이러한 취약점을 개선하기 위해 본 연구에서는 직전 2 프레임에서의 음성 존재 여부와 음성 스펙트럼 변이를 이용한 가변적인 파라미터 값을 사용하여 주어진 환경에 따라 변화함은 물론, 음성과 음성 사이의 상호 연관성을 고려한 향상된 global soft decision 기법을 제시한다.

먼저 직전 2 프레임이 모두 음성 존재일 경우와 그 이외의 경우를 구분하기 위해 식(7)에 따른 GSAP 값을 이용하여 다음과 같이 정의한다.

$$\begin{aligned} P(H_0|Y(n)) > P_{th} &: H(n) = 0 \\ P(H_0|Y(n)) \leq P_{th} &: H(n) = 1. \end{aligned} \quad (11)$$

여기서 P_{th} ($=0.7$)는 음성 존재를 결정하기 위한 임계값이며, $H(n)$ 는 음성 존재 여부를 나타내는 것으로서 음성 존재일 경우에는 1로 정의하고 음성 부재일 경우에는 0으로 정의하여 다음 식에 의해 직전 2 프레임의 상태를 판단한다.

$$I_j = H(n-1)H(n-2). \quad (12)$$

여기서 I_j 는 인디케이터 값으로서 직전 2 프레임이 모두 음성인 경우 I_1 ($I_j=1$)와 그 이외의 경우 I_0 ($I_j=0$)로 분류하여 준다. 이 두 가지 경우에 대하여 음성 스펙트럼 변이 값의 범위에 따라 가변적인 q 값을 적용하게 된다. 음성 스펙트럼 변이는 이전 프레임 동안 예측된 평균 long-term 파워 스펙트럼과 현재 파워 스펙트럼의 차이이며 그 식은 다음과 같다.^[8]

$$\Delta(n) = \sum_{k=0}^{K-1} (|Y(k,n)|^2 - \bar{E}(k,n)), \quad (13)$$

여기서 $|Y(k,n)|^2$ 은 현재 파워 스펙트럼을, $\bar{E}(k,n)$ 는 전 프레임 동안 예측된 평균 long-term 파워 스펙트럼을 각각 나타낸다. 또한 초기값은 $\bar{E}(k,1) = |Y(k,1)|^2$ 이 되며, $\bar{E}(k,n)$ 는 다음 식에 의해 갱신된다.^[8]

$$\bar{E}(k,n) = \beta \bar{E}(k,n-1) + (1-\beta)|Y(k,n)|^2. \quad (14)$$

위의 식들로부터 얻은 값을 실험을 통해 찾은 최적화된 $\Gamma (= 7000)$ 값 즉, 문턱값에 따라 3가지 조건으로 나누어서 q 값을 추정하게 된다.

$$\begin{aligned} \Delta(n) > \Gamma & : \textit{ascending} \\ -\Gamma < \Delta(n) \leq \Gamma & : \textit{static} \\ \Delta(n) \leq -\Gamma & : \textit{descending}. \end{aligned} \quad (15)$$

먼저 I_1 인 경우에는 현재 프레임도 음성일 확률이 높으므로 $\Delta(n)$ 값이 문턱값을 기준으로 유지되고 있을 때에는 기존의 고정된 값으로 유지시키고, 그 값이 증가하고 있을 때와 감소하고 있을 때에는 아래와 같은 시그모이드 형태 함수를 이용하여 q 값을 점진적으로 높아지게 한다.

$$q = \frac{\delta \exp[-\alpha(\Delta(n) - s_0)]}{1 + \exp(-\alpha(\Delta(n) - s_0))} + \sigma. \quad (16)$$

여기서 상수 δ, σ 값은 다양한 잡음 환경과 여러 SNR에 대해 실험적으로 얻은 최적화된 값으로서 각각 $\delta = 5.3, \sigma = 99$ 으로 적용하였고, 기울기 파라미터 $\alpha = 0.5$, 오프셋 (offset) $s_0 = -9$ 으로 설정하였다.

반대로 I_0 인 경우에는 현재 프레임도 음성이 아닐 확률이 높으므로 $\Delta(n)$ 값이 문턱값을 기준으로 유지되고 있을 때에는 기존의 고정된 값으로 유지시키고, 그 값이 증가하고 있을 때와 감소하고 있을 때에는 q 값을 점진적으로 낮아지게 한다. 이 경우에는 식(16)의 상수 및 파라미터 값들을 각각 $\delta = -4.7, \sigma = 0.9, \alpha = 0.5, s_0 = -9$ 으로 설정하였다.

결과적으로 q 값이 높을 때에는 음성 부재 확률 값을 1에 가깝게 만들어줌으로써 성능을 향상시킨다. 이 두 가지의 경우와 같이 q 값을 적응적으로 증감시키기 위해 적용시킨 시그모이드 형태 함수를 각각 Fig. 1과 Fig. 2에 나타내었다. Fig. 1은 I_1 인 경우, 그리고 Fig. 2는 I_0 인 경우에 $\Delta(n)$ 값에 따라 q 값에 적용된 sigmoid type 함수를 각각 보여 주고 있다. 그리고 Fig. 3은 제안된 스펙트럼 변이 기법을 이용하여 얻은 음성 존재 확률과 시그모이드 함수가 적용되어

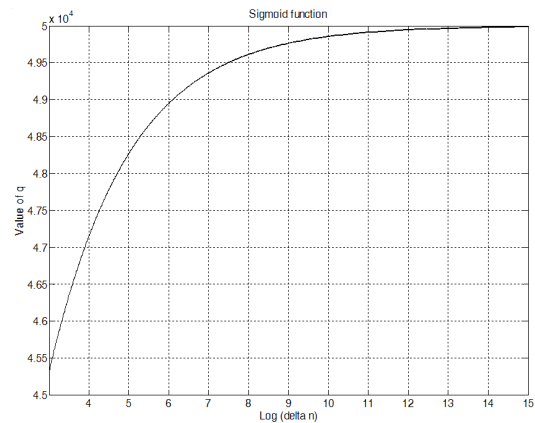


Fig. 1. The sigmoid type function for the value of q (I_1).

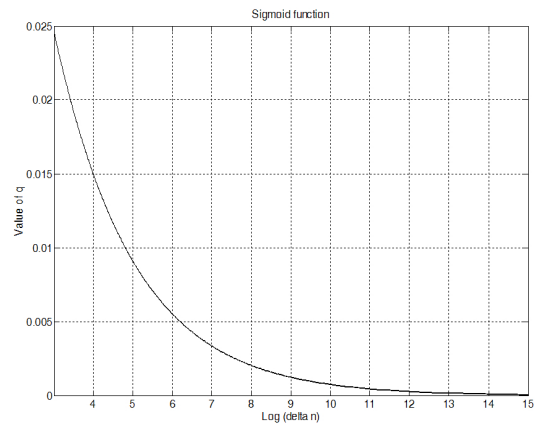


Fig. 2. The sigmoid type function for the value of q (I_0).

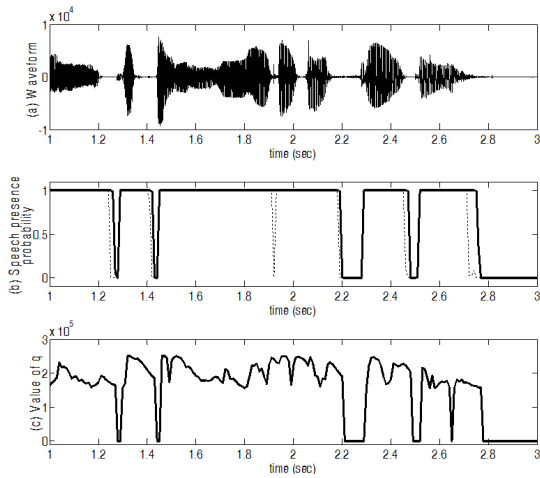


Fig. 3. Comparison of probability under the car noise (SNR=10 dB) (a) Clean speech waveform (b) Speech presence probability in short-time frames: probability of conventional algorithm (dashed line), probability of proposed algorithm (bold line) (c) Proposed value of q .

그에 따라 변화하는 q 값을 보여 주고 있다.

이를 통해 이전의 고정된 파라미터 q 값을 사용하던 global soft decision보다 제안된 스펙트럼 변이 기법을 이용한 방법이 음성 부재 확률을 구할 때 보다 나은 성능을 보임을 확인할 수 있다. 또한 음성이 끝나는 부분에서는 늦게 떨어지면서 음성임에도 불구하고 음성이 아니라고 판단하는 경우를 줄여 주는 것을 볼 수 있다.

IV. 실험 결과

본 연구에서는 기존의 global soft decision 에서 음

성 부재 확률을 구할 때 사용되는 q 값이 고정된 값이었던 것과는 달리,

직전 2 프레임에서의 음성 존재 여부와 스펙트럼 변이 값의 상태 조건에 따라 가변적인 q 값을 적용하여 음성 부재 확률을 향상시키는 기법을 사용하였다. 제안된 알고리즘의 음질 평가를 위해 널리 적용되고 있는 ITU-T P.862 PESQ 방법과^[10] composite measure 중 대표적인 C_{ovl} 방법으로^[11] 음성 향상의 성능 비교를 하였다.

위의 테스트를 위해 한 프레임의 크기가 10ms, 8kHz로 샘플링 된 남성, 여성 화자 각각이 96개의 문장을 발음하도록 한 데이터에 다섯 가지 형태의 잡음이 추가되었다. 잡음은 배블 잡음(babble noise), 자동차 잡음(car noise), 사무실 잡음(office noise), 거리 잡음(street noise), 백색 잡음(white noise)에서 각각 5, 10, 15 dB의 SNR로 추가하였다. 또한 기존의 global soft decision의 고정 파라미터 q 값은 0.0625로 설정해 주었고, 제안된 방법에서 세 가지 조건에서의 각 q 값은 sigmoid type 함수에 의해 얻어진 값이 적용되었다. PESQ 및 C_{ovl} 값은 각각 이들 샘플에 대한 평균 수치로 나타내었고, 전 프레임 동안 예측된 평균 long-term 파워 스펙트럼인 $\bar{E}(k,n)$ 값을 구할 때의 β 값은 0.8로 설정하였다.

Table 1, 2는 기존의 global soft decision 방법보다 논문에서 제안한 스펙트럼 변이 기법을 적용한 것이 PESQ 및 C_{ovl} 수치로 보았을 때, 모든 실험 조건에서 향상된 것을 보여 주고 있다. 즉, 기존의 고정된 파라미터 q 값을 사용하던 global soft decision보다 제안된

Table 1. Comparison of PESQ score in various noise environment.

Noise type	Method	SNR (dB)		
		5	10	15
Babble noise	Conventional	2.258	2.583	2.893
	Proposed	2.264	2.598	2.919
Car noise	Conventional	2.756	3.028	3.354
	Proposed	2.784	3.070	3.392
Office noise	Conventional	2.205	2.528	2.876
	Proposed	2.210	2.538	2.913
Street noise	Conventional	2.648	2.967	3.220
	Proposed	2.687	3.006	3.261
White noise	Conventional	1.989	2.364	2.754
	Proposed	2.007	2.376	2.763

Table 2. Comparison of C_{ovl} score in various noise environment.

Noise type	Method	SNR (dB)		
		5	10	15
Babble noise	Conventional	2.616	3.014	3.368
	Proposed	2.625	3.027	3.394
Car noise	Conventional	3.167	3.490	3.857
	Proposed	3.202	3.539	3.896
Office noise	Conventional	2.613	3.000	3.392
	Proposed	2.621	3.011	3.431
Street noise	Conventional	3.100	3.476	3.756
	Proposed	3.144	3.516	3.798
White noise	Conventional	2.158	2.610	3.050
	Proposed	2.183	2.621	3.058

스펙트럼 변이를 이용한 방법이 다양한 잡음 환경에서 음성 부재 확률을 구할 때, 보다 정확하게 추정할 수 있으므로 음성 향상 시스템에서의 성능이 좋음을 확인할 수 있다.

이는 제안된 알고리즘의 음성 향상 기법이 기존 알고리즘보다 깨끗한 음성 신호에 좀 더 가까운 신호를 생성하는 것을 의미하는 것으로써 음성 부재 확률을 더 잘 추정함에 따라 성능의 향상이 있음을 확인할 수 있었다.

V. 결 론

본 논문에서는 기존의 global soft decision 알고리즘에서 음성 부재 확률의 고정 파라미터 대신 스펙트럼 변이 기법을 이용함으로써, 음성 부재와 존재에 대한 *a priori* 확률값의 비인 q 값을 가변적으로 적용시켜 보다 견실한 음성 부재 확률을 추정하였다.

기존의 global soft decision 방법은 음성 부재 확률을 구하기 위해 기존의 통계적 가정을 바탕으로 고정된 q 값을 적용하였지만, 비정상적인 잡음 환경을 고려하였을 때 잡음이 음성 신호를 간섭하면 음성의 꼬리와 같은 약한 음성 신호를 추정하기 힘들다는 단점을 가지고 있었다.

하지만 제안된 알고리즘에서는 직전 2 프레임에서의 음성 존재 여부와 스펙트럼 변이 값의 상태 조건에 따라 가변적인 q 값을 적용하여 이러한 단점을 극복하였다. 이를 통해 다양한 음성 환경에서의 정확한 음성 부재 확률 추정을 가능하게 하며, 실험 결과 기존의 방법보다 다양한 잡음 환경에 더욱 강인한 성능을 보였다.

감사의 글

이 논문은 2012년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (2012R1A2A2A01004895).

참 고 문 헌

1. S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. on Acoust., Speech, Signal Processing, **27**, 113-120 (1979).
2. J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," IEEE Trans. on Acoust., Speech, Signal Processing, **67**, 1583-1604 (1979).
3. R. J. McAulary and M. L. Malpass, "Speech enhancement using a soft-decision noise suppression filter," IEEE Trans. on Acoust., Speech, Signal Processing, **28**, 137-145 (1980).
4. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," IEEE Trans. on Acoust., Speech, Signal Processing, **32**, 1109-1121 (1984).
5. N. S. Kim and J.-H. Chang, "Spectral enhancement based on global soft decision," IEEE Signal Processing Letters, **7**, 108-110 (2000).
6. D. Malah, R. Cox and A.J. Accardi, "Tracking speech-presence uncertainty to improve speech enhancement in nonstationary noise environments," Proc. IEEE Int. Conf. Acoust. Speech Signal Process., 789-792 (1999).
7. W. Lee, J.-H. Song, and J.-H. Chang, "Minima-controlled speech presence uncertainty tracking method for speech enhancement," Signal Processing, **91**, 155-161 (2011).
8. S.-K. Kim and J.-H. Chang, "Voice activity detection based on conditional MAP criterion incorporating the spectral gradient," Signal Processing, **92**, 1699-1705 (2012).
9. J.-M. Kum and J.-H. Chang, "Improved global soft decision incorporating second-order conditional MAP in speech enhancement," IEICE Transactions on Information and Systems, **93**, 1652-1655 (2010).
10. ITU-T P.862, *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*, 2001.
11. Y. Hu and P. Loizou, "Evaluation of objective quality measures for speech enhancement," IEEE Trans. Audio Speech Language Process., **16**, 229-238 (2008).

저자 약력

▶ 김 종 웅 (Jong-Woong Kim)



2006년 2월: 인하대학교 전자공학과 학사
2012년 2월~현재: 한양대학교 전자컴퓨터통신공학과 석사과정

▶ 장 준 혁 (Joon-Hyuk Chang)



2004년 2월: 서울대학교 전기컴퓨터공학부 박사
2000년 3월~2005년4월: (주)넷더스 연구소장
2004년 5월~2005년4월: 캘리포니아 주립대학 산타바바라(UCSB) 박사후 연구원
2005년 5월~2005년8월: 한국과학기술연구원(KIST) 연구원
2005년 9월~2011년 2월: 인하대학교 전자공학부 조교수
2011년 3월~현재: 한양대학교 융합전자공학부 부교수