

논문 2008-45SP-4-12

실시간 변별적 가중치 학습에 기반한 음성 검출기

(Voice Activity Detection Based on Real-Time Discriminative Weight Training)

강상익*, 조규행*, 장준혁**

(Sang-Ick Chang, Q-Haing Jo, and Joon-Hyuk Chang)

요약

본 논문에서는 다양한 잡음 환경에서 음성의 통계적 모델에 기반한 음성 검출기의 성능향상을 위해 PSFM (Power Spectral Flatness Measure)을 이용하여 실시간으로 변별적 가중치 학습 (Discriminative Weight Training) 기반의 최적화된 우도비 테스트 (Likelihood Ratio Test, LRT)를 제안한다. 먼저, 기존의 통계모델기반의 음성 검출기를 분석하고, 이를 기반으로 MCE (Minimum Classification Error) 방법을 도입하여 도출한 각 주파수 채널별 가중치를 PSFM 값에 기반하여 실시간 매 프레임마다 다른 가중치를 적용한 우도비 기반의 음성 검출 결정법을 제시한다. 제안된 알고리즘은 다양한 잡음 환경에서 기존에 제시된 음성 검출기와 비교하였으며, 우수한 성능을 보인다.

Abstract

In this paper, we apply a discriminative weight training employing power spectral flatness measure (PSFM) to a statistical model-based voice activity detection (VAD) in various noise environments. In our approach, the VAD decision rule is expressed as the geometric mean of optimally weighted likelihood ratio test (LRT) based on a minimum classification error (MCE) method which is different from the previous works in that different weights are assigned to each frequency bin and noise environments depending on PSFM. According to the experimental results, the proposed approach is found to be effective for the statistical model-based VAD using the LRT.

Keywords : 음성 검출기, Minimum Classification Error (MCE), 통계적 모델, 우도비, Power Spectral Flatness Measure (PSFM)

I. 서 론

음성과 비음성 구간을 검출하는 음성 검출기 (voice activity detector, VAD)는 음성 부호화, 음성인식 그리고 음향학적 반향제거기 등 음성 통신 시스템에서 많이 적용된다. 특히, 음성 검출기는 다중 접속 기술에서 한정된 주파수 대역을 효율적으로 사용하기 위한 가변 전송률 부호화기의 실현을 위해 필수적인 부분을 차지하고 있으며 이와 관련하여 다양한 형태의 알고리즘이 제안되고 있다. 기존에 많이 알려진 음성 검출 알고리즘

으로는 선형 예측 부호화 (linear prediction coding, LPC) parameters^[1], 에너지 레벨, 포만트 (formant) 모양^[2], 영교차율 (zero crossing rate, ZCR)^[3], 켭스트럴 피쳐 (cepstral feature)^[4], 그리고 주기적인 척도 (periodicity measure)^[5] 등이 있는데 이러한 많은 알고리즘 중에 Ephraim과 Malah의 연구에서 시작된 minimum mean square error (MMSE) 기반의 음성향상 기법에 사용된 음성의 존재와 부재에 대한 통계적 모델을 음성 검출기에 적용한 것이 매우 우수한 성능을 가진 것으로 알려져 있다^[6~13]. 구체적으로 음성에 대한 가우시안 통계모델을 decision-directed (DD) 기법에 도입하여 최적의 음성 검출 파라미터 추정에 사용하여 나온 음성의 존재와 부재에 대한 우도비 (likelihood ratio,

* 학생회원, ** 정회원, 인하대학교 전자공학부
(School of Electronic Engineering, Inha University)
접수일자: 2007년11월6일, 수정완료일: 2008년4월21일

LR)를 기하 평균한 결정식으로 음성 검출여부를 판단 한다^[8].

본 논문에서는 기존의 음성의 통계적 모델 기반의 음성 검출기에서 제시된 각 주파수 채널별 우도비의 단순한 기하 평균을 이용하여 문턱값을 비교하는 방법 대신, 변별적 가중치 학습 (discriminative weight training)을 위한 minimum classification error (MCE) 방법을 이용하여 도출된 최적화된 가중치를 power spectral flatness measure (PSFM) 값에 따라 실시간으로 매 프레임마다 각 주파수 채널별 우도비에 다르게 적용하여 기하 평균을 구성하는 새로운 방식을 제안하며, 다양한 잡음 환경에서 기존의 통계적 모델 기반의 음성 검출기와 성능을 비교하였다.

본 논문의 II장에서는 통계적 음성 검출기에 대해 소개하고, III장에서는 MCE 훈련을 이용한 각 잡음 환경에 최적화된 가중치를 PSFM 값에 따라 적용한 새로운 음성 검출기에 대하여 논하였다. IV장에서는 기존의 방식과 비교실험 결과를 보여주며, 마지막으로 V장에서 결론을 맺어 본 논문을 마친다.

II. 통계적 음성 검출기

시간축 상에서 원래의 음성신호 $x(t)$ 에 잡음 신호 $n(t)$ 이 인가된 입력신호 $y(t)$ 를 discrete Fourier transform (DFT)을 통해 주파수 축으로 변환되어 아래와 같이 표현된다.

$$Y(t) = X(t) + N(t) \quad (1)$$

여기서 $Y(t) = [Y_1, Y_2, \dots, Y_M]$, $X(t) = [X_1, X_2, \dots, X_M]$, 그리고 $N(t) = [N_1, N_2, \dots, N_M]$ 은 각각 잡음에 오염된 음성신호, 원래의 음성신호, 잡음신호의 DFT 계수 벡터를 나타낸다. 주어진 H_0 , H_1 이 각각 음성의 부재와 존재를 표현한다고 하면 각 주파수 채널별로 다음과 같이 기술된다.

$$H_0: \text{speech absent} : Y_k(t) = N_k(t) \quad (2)$$

$$H_1: \text{speech present} : Y_k(t) = X_k(t) + N_k(t). \quad (3)$$

음성과 잡음신호의 스펙트럼이 복소 가우시안 분포를 따른다는 가정으로부터 가설 H_0 와 H_1 을 조건으로 한 확률밀도함수는 아래와 같이 주어진다^[8].

$$p(Y_k | H_0) = \frac{1}{\pi \lambda_{d,k}} \exp \left\{ -\frac{|Y_k|^2}{\lambda_{d,k}} \right\} \quad (4)$$

$$p(Y_k | H_1) = \frac{1}{\pi [\lambda_{d,k} + \lambda_{x,k}]} \exp \left\{ -\frac{|Y_k|^2}{\lambda_{d,k} + \lambda_{x,k}} \right\} \quad (5)$$

여기서 $\lambda_{x,k}$ 와 $\lambda_{d,k}$ 는 각각 채널별 음성과 잡음의 분산이며, 이 때 k 번째 주파수 밴드에 대한 우도비는 아래와 같이 구한다.

$$\Lambda_k \equiv \frac{p(Y_k | H_1)}{p(Y_k | H_0)} = \frac{1}{1 + \xi_k} \exp \left\{ \frac{\gamma_k \xi_k}{1 + \xi_k} \right\} \quad (6)$$

여기서 $\xi_k = \lambda_{x,k} / \lambda_{d,k}$ 와 $\gamma_k = Y_k / \lambda_{d,k}$ 는 각각 *a priori* signal-to-noise ratio (SNR)와 *a posteriori* SNR이다 [8]. 음성 부재 구간에서 생성되는 잡음 신호로부터 구한 잡음 분산 $\lambda_{d,k}$ 를 이용하여 *a posteriori* SNR γ_k 를 추정하며, 또한 *a priori* SNR ξ_k 는 decision-directed (DD) 방식을 이용하여 아래와 같이 추정한다^[6].

$$\hat{\xi}_k(t) = \alpha \frac{|\hat{X}_k(t-1)|^2}{\lambda_{d,k}(t-1)} + (1-\alpha) P[\gamma_k(t)-1] \quad (7)$$

여기서 $|\hat{X}_k(t-1)|^2$ 은 이전 프레임에서 추정된 음성 신호의 k 번째 스펙트럼 성분의 크기에 대한 추정치이며, MMSE에 기반하여 구한다^[8]. 또한 α 는 가중치 값이며, 연산자 $P[\cdot]$ 은 아래와 같이 정의된다.

$$P[x] = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

기존의 일반적인 통계적 모델 기반의 음성 검출기에 대한 결정식은 각각의 주파수 채널에서 구해진 우도비를 기하 평균하여 아래와 같이 음성 검출 여부를 판단한다^[7~13].

$$\log \Lambda(t) = \frac{1}{M} \sum_{k=1}^M \log \Lambda_k \stackrel{H_1}{>} \stackrel{H_0}{<} \eta \quad (9)$$

여기서 M 은 전체 주파수 대역의 개수이며, η 는 음성 검출 문턱값이다.

III. 다양한 잡음 환경에 따른 최적화 된 가중치 적용

기존의 음성의 통계모델에 기반한 음성 검출기는 식 (9)에서 보듯이 각 주파수별 성분이 독립이라는 가정에 기반하여 간단한 우도비의 기하 평균식을 이용한다. 그러나, 각 주파수 채널별 우도비가 서로 다른 잡음 환경에서 음성 검출 성능에 균일한 기여를 한다는 것은 음성신호의 주파수특성의 분포 등을 고려하면 실제적이지 않다. 따라서, 본 논문에서는 각 우도비 $\log A_k$ 에 각 잡음 환경에 최적화된 가중치를 인가함으로써 보다 효과적인 음성 검출기를 제안하고 새로운 결정식을 다음과 같이 정의한다.

$$A_w = \frac{1}{M} \sum_{k=1}^M w_k \log A_k > \eta \quad (10)$$

여기서 w_k 는 변별적 가중치 학습에 기반한 최적화 된 가중치이다.

1. MCE 훈련을 이용한 다양한 잡음 환경에 최적화된 가중치의 도출

입력 신호로부터 구한 각 주파수 채널별 우도비에 각 다른 가중치 w_k 를 적용하여 새로운 우도비 $w_k \log A_k$ 를 구하며 각 가중치는 다음의 조건을 만족한다.

$$\sum_{k=1}^M w_k = 1, \quad w_k \geq 0 \text{ for } k = 1, \dots, M. \quad (11)$$

최적화된 가중치를 적용한 우도비 벡터를 $A_w = \{w_1 \log A_1, w_2 \log A_2, \dots, w_M \log A_M\}$ 라 하고 A_w 를 $\frac{1}{M} \sum_{k=1}^M w_k \log A_k$ 와 같이 정의한다. 그리고 훈련할 데이터의 각각의 프레임에서 음성 $g_s(\cdot)$ 와 비음성 $g_n(\cdot)$ 을 구분하는 두 개의 함수를 다음과 같이 정의 한다.

$$g_s(A_w) = A_w - \theta \quad (12)$$

$$g_n(A_w) = \theta - A_w \quad (13)$$

여기서 θ 는 음성과 비음성을 구분하는 문턱값이다. 이

때 문턱값은 음성과 비음성 훈련 데이터의 분포에서 얹치는 경계값 ($\theta = 0$)을 사용하였다. 제안된 연구에서는 최적화 알고리즘에 기반한 가중치를 구하기 위해 generalized probabilistic descent (GPD) 기반의 MCE 훈련을 적용하며^[14], 실제로 훈련 데이터 프레임 $A_w(t)$ 의 분류 오류 D 를 다음과 같이 정의한다.

$$D(A_w(t)) = \begin{cases} -g_s(A_w(t)) + g_n(A_w(t)) & \text{if } g_s \text{ is true class} \\ -g_n(A_w(t)) + g_s(A_w(t)) & \text{if } g_n \text{ is true class} \end{cases} \quad (14)$$

여기서 식 (14)이 음수인 값을 가질 때 올바른 분류가 되며 이를 기반으로 하는 손실함수 (loss function) L 은 다음과 같이 sigmoid 함수 형태로 정의된다.

$$L = \frac{1}{1 + \exp(-\beta D(A_w))}, \quad \beta > 0 \quad (15)$$

여기서 β 는 sigmoid 함수의 기울기를 나타낸다. 최적화 된 가중치를 구하기 위해선 손실함수가 최소가 되어야 한다. MCE 훈련과정을 통해 가중치를 조정하는 과정에서 식 (10)과 같은 제약조건 때문에 가중치 w 를 \tilde{w} 로 변환 한다.

$$\tilde{w} = \{\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_k\} \quad (16)$$

$$\tilde{w}_k = \log w_k. \quad (17)$$

가중치 \tilde{w}_k 는 매 프레임마다 연속적으로 존재하는데, 각 주파수 가중치는 다음과 같은 식으로 갱신된다 [15].

$$\tilde{w}_k(n+1) = \tilde{w}_k(n) - \epsilon \frac{\partial L}{\partial \tilde{w}_k} \Big|_{\tilde{w}_k = \tilde{w}_k(n)} \quad (18)$$

여기서 $\epsilon (> 0)$ 는 단조롭게 감소하는 구간의 크기이다. \tilde{w}_k 를 갱신한 후에 아래의 식과 같이 w_k 로 복원된다.

$$w_k = \frac{\exp(\tilde{w}_k)}{\sum_{i=1}^M \exp(\tilde{w}_i)} \quad (19)$$

식 (19)에서 정규화 된 가중치를 사용했을 때 식 (11)을 만족한다.

2. PSFM 기반의 실시간 가중치 구현

다양한 잡음 환경에서 보다 효과적인 각 주파수 채널별 가중치 적용을 위해 본 논문에서는 음성 신호의 매 프레임 마다 서로 다른 채널별 가중치를 선택하기 위해

PSFM 방법을 도입한다^[16]. 기본적으로 PSFM은 다음과 같이 정의되며

$$\Xi(t) = \frac{\sqrt{M} \prod_{k=0}^{M-1} \hat{\lambda}_{x,k}(t)}{\frac{1}{M} \sum_{k=0}^{M-1} \hat{\lambda}_{x,k}(t)} \quad (20)$$

여기서 PSFM $\Xi(t)$ 는 서로 다른 주파수 구간에서의 전력 추정치의 기하 평균 $\sqrt{M} \prod_{k=0}^{M-1} \hat{\lambda}_{x,k}(t)$ 대 산술 평균 $\frac{1}{M} \sum_{k=0}^{M-1} \hat{\lambda}_{x,k}(t)$ 비로 나타낸다. PSFM은 0과 1사이의 값이며 또한 입력 신호의 전력 스펙트럼 추정치는 다음과 같이 주어진다.

$$\hat{\lambda}_{x,k}(t) = \alpha_x \hat{\lambda}_{x,k}(t-1) + (1 - \alpha_x) |X_k(t)|^2 \quad (21)$$

여기서 α_x 는 평활 계수 (smoothing coefficient)이며, 실험적으로 0.5로 설정하였다.

$$\Xi_x(t) = \alpha_{\Xi} \Xi_x(t-1) + (1 - \alpha_{\Xi}) \frac{\sqrt{M} \prod_{k=0}^{M-1} \hat{\lambda}_{x,k}(t)}{\frac{1}{M} \sum_{k=0}^{M-1} \hat{\lambda}_{x,k}(t)} \quad (22)$$

여기서 α_{Ξ} 는 평활 계수이며, 실험적으로 0.5로 설정하였고 $\Xi_x(t)$ 는 프레임 간 상관성을 고려한 smoothed PSFM이다. 그림 1은 다양한 잡음 환경 및 다양한 SNR에서 음성 신호를 MCE 훈련 방법을 통해 도출된 가중치 분포이다. 구체적으로 128 포인트크기의 DFT를 채택했을 때 얻어지는 63개의 주파수 계수를 16개의 서브대역으로 결합하여 실험한 결과이다. 실제로, PSFM 분포에 따라 두 개의 구간으로 나누었으며 가중치 w_k 는 다음과 같이 정의한다.

$$w_k = \begin{cases} w_k^A & 0 \leq \Xi < \Xi_0 \\ w_k^B & \Xi_0 \leq \Xi < 1 \end{cases} \quad (23)$$

여기서 Ξ_0 는 실험적으로 0.44로 결정한 문턱값이며 w_k^A , w_k^B 는 그림 1에서 각각 O, X으로 나타낸 다른 가중치이다.

기준의 통계적 모델 기반의 음성 검출기의 결정식인 식 (9)와 비교하여 본 논문에서는 위에서 제시된 각 잡음 환경에서 MCE 훈련방법을 이용해 구한 식 (19)의

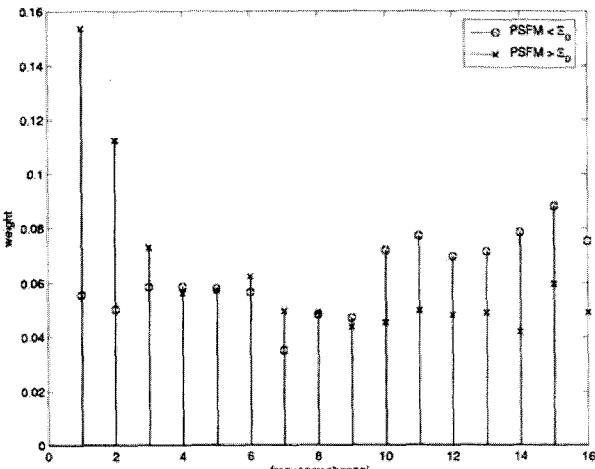


그림 1. PSFM에 따른 가중치 분포

Fig. 1. Weights distribution according to PSFM.

가중치인 w_k^A 와 w_k^B 를 입력신호의 PSFM 값에 따라 매 프레임마다 다르게 각각의 주파수 채널에 곱해서 구해진 우도비를 기하 평균하여 식 (10)과 같이 최종적으로 음성 검출 여부를 판단한다.

IV. 실험

본 논문에서 제안된 PSFM에 기반한 최적화된 가중치 기반의 실시간 음성 검출기의 성능을 평가하기 위해 기존의 우도비 테스트를 이용한 통계적 모델 기반의 음성 검출기의 성능과 동작 특성 (receiver operating characteristic, ROC) 곡선을 이용하여 비교하였다^[8]. 최적화된 가중치를 도출하기 위해 손실함수 L 에서 정의된 기울기 파라미터 $\beta = 1$ 로 결정하였다. 실험에 사용된 데이터는 성능 평가를 위해 총 230초의 깨끗한 음성 데이터에 음성과 비음성 부분을 10 ms마다 수동으로 표시하였다. 분류된 음성 데이터의 음성 구간은 총 57.1%로 유성음 44.0%, 무성음 13.1%로 구성되었으며 잡음 환경은 음성 데이터에 street, car, office, white 잡음이 각각 5, 10, 15 dB SNR로 부과되었다. 그림 2, 3은 5, 15 dB SNR에서 각각의 잡음 환경에 대해 독립적으로 문턱값을 변경하면서 실제 음성을 음성이라고 판단한 음성 검출 확률 (P_d)과 비음성에 대해 음성이라고 판단한 오경보 확률 (P_f)을 측정한 ROC 곡선이다. 그림 4는 각각 5, 10, 15 dB SNR에서 street, car, office, white 잡음 환경이 3~5 s마다 불규칙하게 변할 때 음성 검출기의 문턱값을 변경하면서 실제 음성을 음성이라고 판단한 음성 검출 확률 (P_d)과 비음성에 대해 음성이라고 판단한 오경보 확률 (P_f)

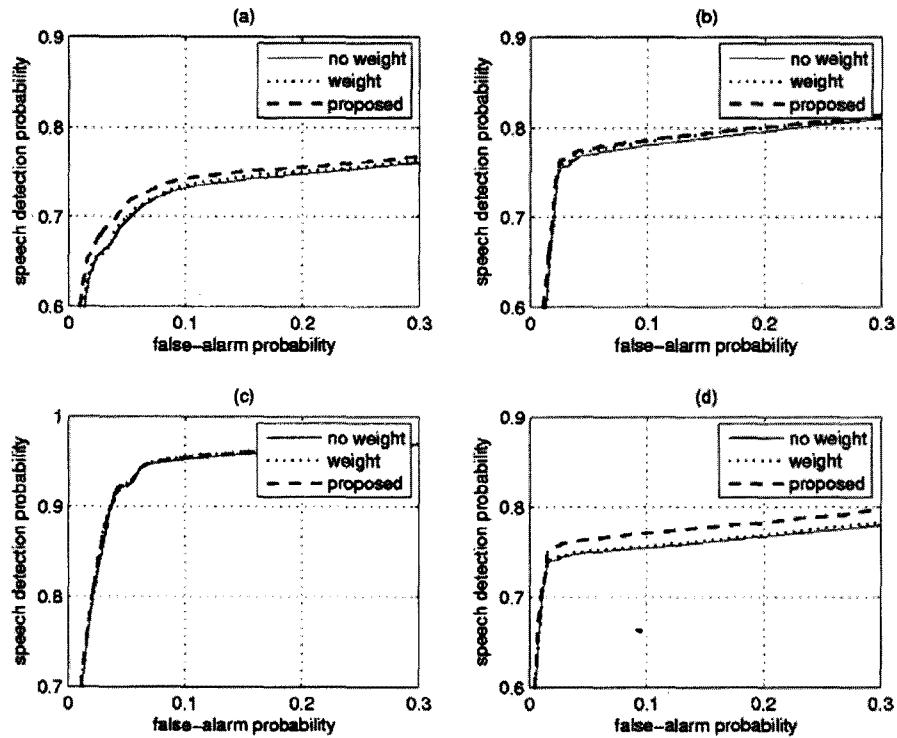


그림 2. 다양한 잡음 환경에서의 ROC 곡선 (5 dB SNR)

(a) street (b) car (c) office (d) white

Fig. 2. ROC curves for various noise environments.

(5 dB SNR) (a) street (b) car (c) office (d) white.

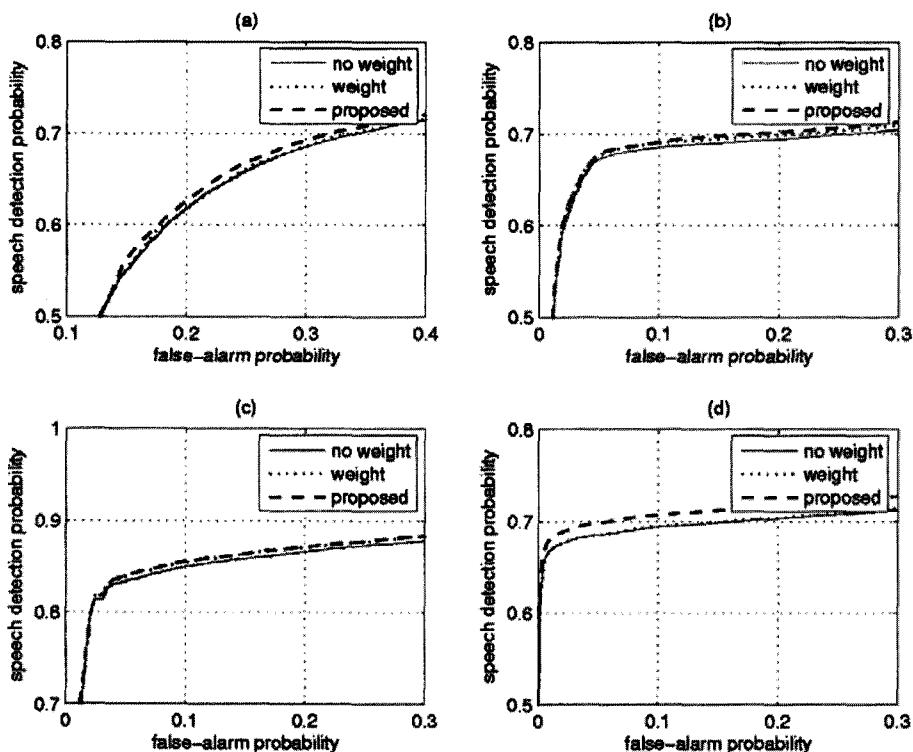


그림 3. 다양한 잡음 환경에서의 ROC 곡선 (15 dB SNR)

(a) street (b) car (c) office (d) white

Fig. 3. ROC curves for various noise environments. (15 dB SNR)

(a) street (b) car (c) office (d) white.

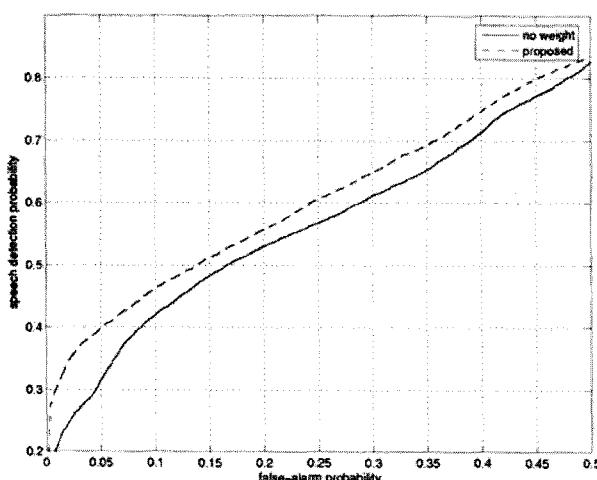


그림 4. 다양한 잡음 환경과 SNR이 불규칙하게 변할 때의 ROC 곡선

Fig. 4. ROC curves for various noise environments and SNRs.

을 측정한 ROC 곡선이다. 이 때 각각의 잡음 환경의 비율은 street, car, office, white 잡음이 각각 25.22%, 25.22%, 23.91%, 25.65%로 구성되어 있고 SNR의 비율은 5, 10, 15 dB가 각각 34.78%, 35.21%, 30%로 구성되어 있다. 실험 분석 결과 동일한 SNR의 주어진 잡음 조건에서 PSFM을 이용하여 각 채널별 잡음 환경에 최적화된 가중치를 적용한 음성 검출기의 경우 전체적으로 기존의 음성 검출기보다 향상된 성능을 보여주며, 특히 동일한 가중치를 적용하였을 때보다 office, white 잡음 환경에서 뚜렷한 성능 개선이 이루어졌다. 뿐만 아니라 그림 4를 통해 잡음 환경과 SNR이 무작위로 변경되는 조건에서도 제안된 방법이 기존의 방법보다 향상된 성능을 볼 수 있다.

결론적으로 기존의 우도비의 기하 평균으로 음성을 검출하는 것보다 PSFM을 이용해 각 채널별 최적화된 가중치를 매 프레임마다 적용하여 음성 검출 방법이 우수한 것을 확인할 수 있다.

V. 결 론

본 논문에서는 음성의 존재와 부재에 대한 통계적 모델에 기반한 각 주파수 채널별 우도비를 단순히 기하 평균을 취하여 문턱값과 비교하는 기존의 방법 대신, MCE 훈련을 이용하여 도출한 최적화된 가중치를 매 프레임마다 PSFM을 이용하여 잡음 환경을 예측하여 각 주파수 채널별 우도비에 적용하는 새로운 방법으로 기존의 방식보다 향상된 통계모델기반의 음성 검출기를

제시하였으며, 객관적인 실험 결과로부터 제안된 음성 검출기의 성능이 우수함을 알 수 있었다.

감사의 글

본 연구는 정보통신연구진흥원의 IT핵심기술개발사업의 일환으로 수행하였음.

참 고 문 헌

- [1] L. R. Rabiner and M. R. Sambur, "Voiced-unvoiced-silence detection using Itakura LPC distance measure," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 323-326, May 1977.
- [2] J. D. Hoyt and H. Wechsler, "Detection of human speech in structured noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 237-240, May 1994.
- [3] J. C. Junqua, B. Reaves, and B. Mark, "A study of endpoint detection algorithms in adverse conditions: Incidence on a DTW and HMM recognizer," in *Proc. Eurospeech*, pp. 1371 - 1374, 1991.
- [4] J. A. Haigh and J. S. Mason, "Robust voice activity detection using cepstral feature," in *Proc. IEEE TELCON*, pp. 321 - 324, China, 1993.
- [5] R. Tucker, "Voice activity detection using a periodicity measure," in *Proc. Inst. Electr. Eng.*, vol. 139, pp. 377 - 380, Aug. 1992.
- [6] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoustics, Speech, Sig. Process.*, vol. ASSP-32, no. 6, pp. 1190-1121, Dec. 1984.
- [7] J. Sohn and W. Sung, "A voice activity detector employing soft decision based noise spectrum adaptation," in *Proc. Int. Conf. Acoustics, Speech, and Sig. Process.*, vol. 1, pp. 365-368, May 1998.
- [8] J. Sohn, N. S. Kim, and W. Sung, "A statistical model-based voice activity detection," *IEEE Sig. Process. Lett.*, vol. 6, no. 1, pp. 1-3, Jan. 1999.
- [9] Y. D. Cho and A. Kondoz, "Analysis and improvement of a statistical model-based voice activity detector," *IEEE Sig. Process. Lett.*, vol. 8, no. 10, pp. 276-278, Oct. 2001.
- [10] J. -H. Chang, J. W. Shin, and N. S. Kim,

- "Voice activity detector employing generalised gaussian distribution," *Electron. Lett.*, vol. 40, no. 24, pp. 1561–1563, Nov. 2004.
- [11] J. -H. Chang, N. S. Kim, and S. K. Mitra, "Voice activity detection based on multiple statistical models," *IEEE Trans. Sig. Process.*, vol. 54, no. 6, pp. 1965–1976, June 2006.
- [12] Y. C. Lee and S. S. Ahn, "Statistical model-based VAD algorithm with wavelet Transform," *IEICE Trans. Fundamentals.*, vol. E89-A, no. 6, pp. 1594–1600, June 2006.
- [13] J. Ramirez, J. M. Gorriz, J. C. Segura, C. G. Puntonet, and A. J. Rubio, "Speech/non-speech discrimination based on contextual information integrated bispectrum LRT," *IEEE Sig. Process. Lett.*, vol. 13, no. 8, pp. 497–500, Aug. 2006.
- [14] B. -H. Juang, W. Chou, and C. -H. Lee, "Minimum classification error rate methods for speech recognition," *IEEE Trans. Speech Audio Processing*, vol. 5, no. 3, pp. 257–265, May 1997.
- [15] Y. Kida and T. Kawahara, "Voice activity detection based on optimally weighted combination of muliple feature," in *Proc. Interspeech*, pp. 2621–2624, Sep. 2005.
- [16] J. -H. Chang, S. Gazor, N. S. Kim, and S. K. Mitra, "Multiple Statistical Models for Soft Decision in Noisy Speech Enhancement," *Pattern Recognition*, vol. 40, no. 3, pp. 1123–1134, Mar. 2007.

저자 소개



강 상 익(학생회원)
2007년 인하대학교 전자공학과
학사 졸업.
2007년 ~ 현재 인하대학교 전자
공학과 석사과정.
<주관심분야 : 음성검출, 음성인
식>



조 규 행(학생회원)
2004년 인하대학교 전자공학과
학사 졸업.
2004년 ~ 2006년 LG.Philips LCD
연구원.
2006년 ~ 현재 인하대학교
전자공학과 석사과정.
<주관심분야 : 음성검출, 잡음제거>



장 준 혁(정회원)
1998년 경북대학교 전자공학과
학사 졸업.
2000년 서울대학교 전기공학부
석사
2004년 서울대학교 전기컴퓨터
공학부 박사
2000년 ~ 2005년 (주)넷더스 연구소장.
2004년 ~ 2005년 캘리포니아 주립대학,
산타바바라 (UCSB) 박사후연구원.
2005년 ~ 2005년 한국과학기술연구원 (KIST)
연구원.
2005년 ~ 현재 인하대학교 전자공학부 조교수.
<주관심분야 : 음성/오디오 신호처리, 통신 신호
처리, 휴면/컴퓨터 인터페이스 등>