

청각장애인을 위한 웨어러블 기기의 위험소리 검출 엔진 설계

A Design of Dangerous Sound Detection Engine of Wearable Device for Hearing Impaired Persons

변성우* · 이석필†

(Sung-Woo Byun · Soek-Pil Lee)

Abstract - Hearing impaired persons are exposed to the danger since they can't be aware of many dangerous situations like fire alarms, car horns and so on. Therefore they need haptic or visual informations when they meet dangerous situations. In this paper, we design a dangerous sound detection engine for hearing impaired. We consider four dangerous indoor situations such as a boiled sound of kettle, a fire alarm, a door bell and a phone ringing. For outdoor, two dangerous situations such as a car horn and a siren of emergency vehicle are considered. For a test, 6 data sets are collected from those six situations, we extract LPC, LPCC and MFCC as feature vectors from the collected data and compare the vectors for feasibility. Finally we design a matching engine using an artificial neural network and perform classification tests. We perform classification tests for 3 times considering the use outdoors and indoors. The test result shows the feasibility for the dangerous sound detection.

Key Words : Hearing impaired, dangerous sound detection engine, LPC, LPCC, MFCC

1. 서론

최근 빠르게 진행되고 있는 고령화 사회로 인해 노인 인구가 늘어나고 난청인구가 증가하고 있으며, 청각장애인의 숫자 또한 통상 인구의 3% 정도까지 크게 증가하고 있다[1]. 청각장애인은 소리를 듣고 상황을 판단하는 것이 어렵기 때문에 일상생활에 많은 어려움이 있을 뿐만 아니라 소리 정보를 이용하여 실내, 실외 환경에서의 위험한 상황을 인지 할 수 없어 즉각적인 대처가 불가능하다. 특히, 위험상황에서 청각장애인들은 시각정보에 의존하게 된다. 따라서 청각장애인들이 위험에 처했을 때, 몸에 부착된 센서 또는 휴대용 기기를 통해 위험상황을 시각, 촉각 정보로 변환하여 전달할 필요가 있다. 이에 따라 청각장애인들에게 위험상황을 빠르게 전달 해줄 수 있는 웨어러블 기기의 연구와 개발이 필요하다.

청각장애인들을 위한 위험상황 감지 기술에 대한 연구에는 MFCC (Mel Frequency Cepstral Coefficient) 특징벡터 기반의 GMM (Gaussian Mixture Model) 패턴인식 알고리즘을 이용하여 오토바이, 차 경적, 천둥, 대형자동차, 공사장, UBM(Universal Background Model), 총 6가지의 상황 인식을 위한 연구[1], Mel-Frequency Cepstral Coefficient(MFCC)를 기반의 Hidden Markov Model(HMM)을 사용한 연구[2], 음향 간의 스펙트럼 특

징의 차이를 이용한 연구[3] 등이 있다.

본 논문에서는 ANN(Artificial Neural Network)을 이용하여 청각장애인을 위한 웨어러블 기기의 위험상황 검출 엔진을 설계한다. 이를 위하여 위험상황을 실내 위험상황, 실외 위험상황으로 나누고 실내 위험상황에는 물 끓는 소리, 화재경보, 초인종, 전화벨로 정의하였고 실외 위험상황에는 사이렌, 차 경적 소리로 정의하였다. 실험데이터는 각 6개의 위험상황별 데이터를 수집하였고 검출 엔진 실험을 위해 환경잡음을 조용한방, 골목길, 번화가, 차도 총 4가지로 정의하고 약 5분가량의 데이터를 수집하였다. 수집된 위험상황별 데이터를 환경잡음과 합성하였고 신호처리 분야에서 가장 많이 쓰이는 특징벡터인 LPC, LPCC, MFCC를[4][5][6][7] 사용하여 위험상황 검출 정확도를 비교 분석하였다.

본 논문의 구성은 다음과 같다. 2장에서는 실험 데이터에 대한 설명, 3장에서는 특징벡터에 대한 설명과 각 특징벡터의 분리도 분석 결과를 보여주고 4장에서는 실험 구성 및 결과를 보여주고 5장에서는 결과를 통해 결론을 맺도록 한다.

2. 실험 데이터

2.1 환경잡음 DB

청각장애인들에게 위험상황이 발생할 수 있는 환경은 실내 환경과 실외 환경이 있으며, 두 가지 환경에서 발생할 수 있는 잡음이 다르기 때문에 실내 환경과 실외 환경은 별도의 실험이 필요하다. 따라서 본 논문에서는 실내 환경으로는 조용한 방으로

† Corresponding Author : Dept. of Media Software, SangMyung University, Korea.

E-mail: esprit@smu.ac.kr

* Dept. of Computer Science, SangMyung University, Korea.
Received : June 2, 2016; Accepted : June 27, 2016

정의하였으며 실외 환경으로는 위험 상황이 발생할 수 있는 위치를 고려하여 한적한 골목길, 변화가, 차도 3가지로 정의하였다. 총 4가지 환경잡음은 각각 5분가량 녹음하였으며 16kHz로 샘플링하고 모노 채널을 사용하였다. 녹음 장비로는 실제 보조 장치의 음질을 고려하여 일반적인 휴대용 마이크를 이용하였다.

2.2 위험상황 소리 DB

청각장애인들의 주변 환경에 따라 발생할 수 있는 위험상황 소리가 다르다. 실내 환경에서는 주전자 물 끓는 소리, 화재경보, 전화벨소리, 초인종 등이 있고, 실외 환경으로는 차 경적소리, 응급차량의 사이렌 소리 등이 있다. 따라서 본 논문에서는 실내 환경에서 발생하는 위험 소리는 주전자 물 끓는 소리, 화재경보, 전화벨소리, 초인종 총 4가지로 정의하였고 실외 환경에서 발생하는 위험 소리는 차 경적소리, 응급차량의 사이렌 소리 총 2가지로 정의하였다. 각 위험상황 소리 데이터는 30초 ~ 60초가량 녹음하였고, 16kHz로 샘플링하고 모노 채널을 사용하였다.

3. 특징벡터

MFCC(Mel-Frequency Cepstrum Coefficient)는 frame내의 음성 신호에 대하여 계산한 파워 스펙트럼을 청각기의 주파수 반응도를 모사한 Mel-scale 주파수 도메인에서 DCT(Discrete Cosine Transform)를 취한 값이다. 일반적으로 이를 추출하기 위한 자세한 과정은 아래 순서도와 같다[8].

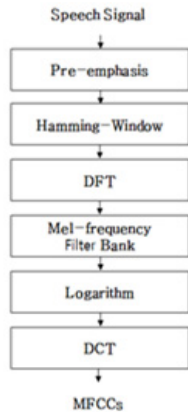


그림 1 MFCC 추출 순서도

Fig. 1 Flow chart of MFCC extraction

LPC는 선형 결합에 의해 과거의 신호에서 현재의 신호[n]을 예측하는 방법으로 전극(All-pole) 모델을 사용하여 식(1)과 같이 차분 방정식의 형태로 나타낼 수 있다. 여기서 S_n 은 입력신호, \tilde{S}_n 은 예측신호, a_i 는 선형예측계수이며, p는 예측계수의 차수이다. 현재신호와 예측된 신호의 예측오차는 식 (2)와 같다.

$$\tilde{S}_n = -(a_1 S_{n-1} + a_2 S_{n-2} + \dots + a_p S_{n-p}) \quad (1)$$

$$e_n = S_n - \tilde{S}_n \quad (2)$$

식 (3)은 예측신호에 대한 mean square error(MSE) J이며, 예측신호의 에러를 최소로 하는 선형예측계수를 찾기 위하여 식(3)을 a_i 에 대해 편미분하면 p개의 선형 연립방정식 (4)을 얻을 수 있다. 식(4)에서 $E[s(n-i)s(n-j)]$ 는 입력신호 S_n 의 자기상관함수이기 때문에 식(4)는 식(5)와 같이 나타낼 수 있고, 선형예측계수는 식(5)에서의 자기 상관 행렬의 역행렬을 이용하여 구할 수 있다. 자기상관함수의 역행렬을 구하는 것의 시간 복잡도 문제를 해결하기 위해 일반적으로 잘 알려져 있는 Levinson-Durbin 알고리즘을 사용하였다 [8].

$$J = E[e^2(n)] = E[(S_n - \tilde{S}_n)^2] \quad (3)$$

$$\sum_{j=1}^p a_j E[s(n-i)s(n-j)] \quad \text{for } i = 1 \dots p \quad (4)$$

$$\begin{bmatrix} R_0 & \dots & R_{p-1} \\ \vdots & \ddots & \vdots \\ R_{p-1} & \dots & R_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_p \end{bmatrix} \quad (5)$$

LPCC는 C(z)의 inverse z-transform으로 정의되고 다음 식(6)과 같다.

$$C(z) = \sum_n c(n)z^{-n} \quad (6)$$

전극(All-pole) $z = z_i$ 가 unit cycle 안에 있고, gain값을 1로 주면 LPCC ($c_{ip}(n)$)는 다음 식(7)과 같이 정의 된다.

$$c_{ip}(n) = \begin{cases} \frac{1}{n} \sum_{i=1}^p z_i^n & n > 0 \\ 0 & n < 0 \end{cases} \quad (7)$$

LPCC는 recursive에 의해 선형예측계수로부터 구한다[8]. recursive 과정은 다음 식 (8)과 같다.

$$\begin{aligned} c_1 &= a_1 \\ c_n &= \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) a_k c_{n-k} + a_n \quad 1 < n < p \\ c_n &= \sum_{k=1}^{n-1} \frac{k}{n} c_k a_{n-k} \quad n > p \end{aligned} \quad (8)$$

3.1 분리도 비교

본 논문에서 쓰이는 MFCC, LPC, LPCC 3가지의 특징벡터들의 분리도를 비교하기 위하여 Bhattacharyya거리측정을 이용하였다.

Bhattacharyya거리 측정은 오류율을 측정하여 거리를 계산하는 방법으로 각 클래스의 분포가 가우시안 형태를 가질 때 가장 좋은 평가기준이 된다. Bhattacharyya거리의 식은 다음 식(9)과 같다. 식에서 M_1, M_2 : 클래스 1,2의 평균, \sum_1, \sum_2 : 클래스 1,2의 공분산이다[9].

$$\mu(1/2) = \frac{1}{8}(M_2 - M_1)^T \left\{ \frac{\sum_1 + \sum_2}{2} \right\}^{-1} (M_1 - M_2) + \frac{1}{2} \ln \frac{|\sum_1 + \sum_2|}{\sqrt{|\sum_1| |\sum_2|}} \quad (9)$$

Bhattacharyya거리가 가장 큰 값이 나온 특징벡터가 클래스 간의 거리가 가장 멀리 떨어져있다는 의미로, 위험상황 검출에 가장 적합하다고 할 수 있다. 총 6가지의 위험상황 소리 실험데이터에서 각 15초가량의 소리를 윈도우 크기 150ms 단위로 특징벡터를 추출하여 각 120개 씩 총 720개를 추출하였다. 이렇게 추출된 특징벡터들의 분리도를 Bhattacharyya거리를 구하여 비교하였다.

표 1, 표 2, 표 3은 각각 6개의 위험상황 소리에서 LPC 특징벡터, LPCC 특징벡터, MFCC 특징벡터를 추출한 후에 각 클래스 간의 거리를 Bhattacharyya거리로 나타낸 표이다. 각 위험상황 소리마다 다른 위험상황과의 거리를 비교 분석 하였다. LPC 특징벡터의 평균 분리도는 2.79로 나왔으며 MFCC 특징벡터의 평균분리도는 3.77로 나왔다. 가장 높게 나온 LPCC 특징벡터의 평균 분리도는 3.79로 나왔다. 이는 LPCC 특징벡터를 사용하였을

표 1 LPC 특징벡터 분리도

Table 1 Separability of LPC feature vector

	물 끓는 소리	사이렌	전화벨	차 경적	초인종	화재 경보
물 끓는 소리	0	2.96	2.63	4.6	2.66	2
사이렌	2.96	0	1.87	2.17	2.51	2.79
전화벨	2.63	1.87	0	2.07	2.27	2.75
차 경적	4.6	2.17	2.07	0	4.27	5.06
초인종	2.66	2.51	2.27	4.27	0	1.25
화재 경보	2	2.79	2.75	5.06	1.25	0

표 2 LPCC 특징벡터 분리도

Table 2 Separability of LPCC feature vector

	물 끓는 소리	사이렌	전화벨	차 경적	초인종	화재 경보
물 끓는 소리	0	2.95	4.89	3.35	4.65	3.21
사이렌	2.95	0	6.06	2.82	5.08	4.44
전화벨	4.89	6.06	0	4.82	2.27	2.77
차 경적	3.35	2.82	4.82	0	3.81	3.49
초인종	4.65	5.08	2.27	3.81	0	2.22
화재 경보	3.21	4.44	2.77	3.49	2.22	0

때, 평균적으로 각 위험상황 소리간의 거리가 가장 멀다는 의미로 통계적으로 LPCC 특징벡터가 각 위험상황의 특징을 잘 반영할 수 있는 특징벡터이다.

표 3 MFCC 특징벡터 분리도

Table 3 Separability of MFCC feature vector

	물 끓는 소리	사이렌	전화벨	차 경적	초인종	화재 경보
물 끓는 소리	0	4.74	6.31	4.68	5.55	5.28
사이렌	4.74	0	5.15	2.33	4.61	4.24
전화벨	6.31	5.15	0	3.11	1.67	2.37
차 경적	4.68	2.33	3.11	0	2.57	2.13
초인종	5.55	4.61	1.67	2.57	0	1.83
화재 경보	5.28	4.24	2.37	2.13	1.83	0

4. 실험 및 결과

그림 5는 위험소리 검출 실험 구성도이다. 실내에서 발생하는 위험상황은 실내의 환경에 비해 즉각적인 알림이 필요하지 않다. 이에 반해 실외에서 발생하는 위험상황은 위험상황을 피하기 위해서 빠르게 알림을 주는 것이 중요하다[10]. 이처럼 실내 환경과 실외 환경에서 처리하는 것이 다르고 발생하는 위험상황이 다르기 때문에 본 논문에서는 실내 환경과 실외 환경 2가지로 나누어서 실험을 진행하였다.

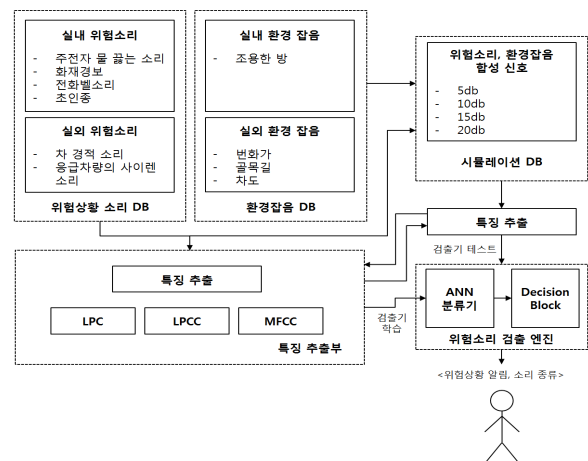


그림 2 위험소리 검출 실험 구성도

Fig. 2 Architecture of dangerous sound detection engine

추출된 특징벡터를 이용하여 위험 상황을 인지하기 위해 ANN 패턴인식 알고리즘을 사용하였다. 신경회로망은 훈련을 통해 각 퍼셉트론간의 연결 강도를 의미하는 가중치 값을 조정하여 훈련

하고 입력의 비선형 함수를 유추할 수 있도록 설계된 패턴 인식 알고리즘이다. 본 연구에서 사용되는 신경회로망은 웨어러블 기기의 컴퓨팅 환경을 고려하여 구조를 복잡하게 하지 않기 위해 그리고 특징벡터를 추출할 때 신호의 성질이 인코딩되기 때문에 input layer와 hidden layer 사이에 신호의 성질을 인코딩할 수 있는 전처리 과정이 들어가지 않는다. 이로 인해 신경회로망 내의 많은 양의 파라미터를 줄일 수 있다. 최근 많은 신경회로망을 사용한 연구에서 영상, 음성 파일을 전처리 과정을 거치지 않고 사용하고 있다[11]. 이와 같이 전체의 영상, 음성 신호를 사용하게 되면 input layer의 크기가 커지게 되고 hidden layer와의 연결 파라미터도 복잡해진다. 컴퓨팅 환경이 좋아지고 연결 파라미터를 줄이기 위한 많은 연구들이 진행되었지만 웨어러블 기기와 같은 임베디드 환경에서 많은 양의 파라미터는 오버피팅을 일으킨다.

본 연구에서 사용하는 신경회로망 구조는 input layer와 hidden layer, output layer로 구성되어 있으며 임베디드 환경을 고려하여 hidden layer는 한 개의 층으로 복잡하지 않게 구성하였으며, input layer의 입력 뉴런의 개수는 추출한 특징벡터의 차수와 같이 10개로 정하였다. hidden layer의 뉴런 개수는 6개로 여러 번 반복하여 실험적으로 결정하였다. 6개의 뉴런 중 하나의 뉴런은 bias 값을 갖는다. output layer는 6개의 뉴런으로 구성되어 있으며 각 뉴런은 6개의 위험상황에 대하여 각각의 클래스 점수를 출력하게 된다. 신경회로망내의 전체 뉴런의 계수는 학습 과정에서 output layer의 출력의 에러에 따라 오차역전파법(Back-Propagation)을 사용하여 에러를 최소화하는 계수를 업데이트 하였다. 마지막으로, 신경회로망에서 사용된 활성화 함수는 $\alpha = 1$ 인 양극 시그모이드를 사용하였고, 계수의 학습 율은 0.6으로 하여 학습을 진행하였다.

검출엔진을 학습시키기 위해 실내, 실외 총 6가지의 위험소리에 대한 실험데이터에서 각 15초가량의 소리를 윈도우 크기 150ms 단위로 특징벡터를 추출하여 각 120개 씩 총 720개를 추출하였다. 이 중 576개는 훈련용으로 나머지 144개는 테스트용으로 사용하였다. LPC, LPCC, MFCC에 대하여 각각 훈련을 진행하였으며, 훈련 정확도는 LPC 특징벡터는 99.86%, LPCC 특징벡터는 99.86% 마지막으로 MFCC 특징벡터는 94.3%의 훈련 정확도를 보였다.

위험소리 검출 실험데이터로는 6가지의 위험상황마다 20초~30초의 데이터를 사용하였으며, 실험을 위해 실내 환경에서 발생하는 주전자 물 끓는 소리, 화재경보, 전화벨소리, 초인종소리는 조용한 방에서 녹음한 실내 환경잡음과 5db, 10db, 15db, 20db로 합성하였다. 실외 환경에서 발생하는 차 경적소리, 응급차량의 사이렌 소리는 변화가, 골목길, 차도 3가지 환경잡음과 각각 5db, 10db, 15db, 20db로 합성하였다. 위험소리와 환경잡음을 합성한 실험데이터에서 LPC, LPCC, MFCC를 추출하고 각각 ANN 분류기를 사용하여 패턴분류를 진행하였다. 약 1초가량의 분류 결과 중 66%이상 이 같은 출력 값을 가질 때 해당 부류로 분류해주는 과정을 거치게 되고 어느 부류도 66%이상 넘지 못하는 경우 없음(none)으로 분류 하였다.

4.1 실내 환경 실험 결과

실내 환경 실험은 실내 환경에서 발생할 수 있는 위험소리와 조용한 방의 환경잡음을 합성하여 상황별 위험소리 검출을 실험하였다.

표 4 LPC 특징벡터를 이용한 실험 결과

Table 4 Accuracy result for indoor situation using LPC feature vector

	5db	10db	15db	20db
물 끓는 소리	77.27	86.36	95.45	95.45
전화벨	60	63.33	73.33	80
초인종	56.67	60	70	76.67
화재 경보	63.16	89.47	94.74	94.74

표 4는 LPC 특징벡터를 사용하여 상황별 위험소리 검출한 실험 결과이다. 5db에서는 평균 64.27%의 검출 율을 보였으며, 10db에서는 74.8%, 15db에서는 83.38%, 20db에서는 86.71%의 검출 율을 보였다. 20db에서 주전자 물 끓는 소리의 경우 최대 95.45%의 검출 율을 보였으며, 초인종 소리의 경우 76.67%로 비교적 검출 율이 낮게 나왔다.

표 5 LPCC 특징벡터를 이용한 실험 결과

Table 5 Accuracy result for indoor situation using LPCC feature vector

	5db	10db	15db	20db
물 끓는 소리	100	100	100	100
전화벨	36.67	53.33	80	80
초인종	44.74	94.74	100	97.34
화재 경보	57.89	63.16	94.74	94.74

표 5는 LPCC 특징벡터를 사용하여 상황별 위험소리 검출한 실험 결과이다. 5db에서는 평균 59.82%의 검출 율을 보였으며, 10db에서는 77.80%, 15db에서는 93.68%, 20db에서는 93.02%의 검출 율을 보였다. 주전자 물 끓는 소리의 경우 5db에서 20db까지 100%의 검출 율을 보였으며, 전화벨소리의 경우 20db에서 80%, 5db에서 36.67%로 비교적 검출 율이 낮게 나왔다.

표 6 MFCC 특징벡터를 이용한 실험 결과

Table 6 Accuracy result for indoor situation using MFCC feature vector

	5db	10db	15db	20db
물 끓는 소리	81.82	86.36	86.36	81.82
전화벨	26.67	73.34	76.67	80
초인종	44.74	78.95	100	100
화재 경보	52.63	89.47	94.74	94.74

표6은 MFCC 특징벡터를 사용하여 상황별 위험소리 검출한 실험 결과이다. 5db에서는 평균 51.46%의 검출 율을 보였으며, 10db에서는 82.03%, 15db에서는 89.44%, 20db에서는 89.13%의 검출 율을 보였다. 초인종 소리의 경우 20db에서 최대 100%의 검출 율을 보였으며, 전화벨소리의 경우 20db에서 80%, 5db에서 26.7%로 비교적 검출 율이 낮게 나왔다.

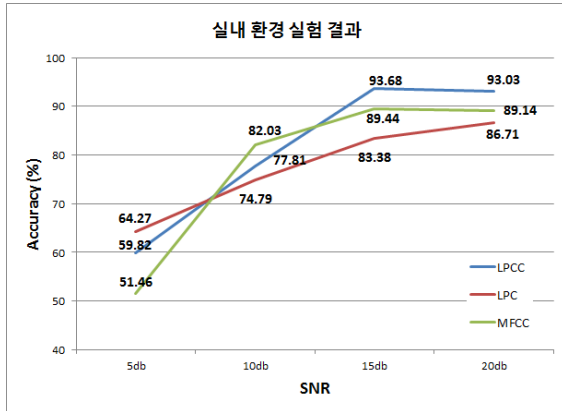


그림 3 실내 환경 실험 결과

Fig. 3 Accuracy comparative experiment result of various SNR for indoor situation

그림 3은 실내 환경 실험에 대한 결과이다. 5db에서 64.27%로 LPC 특징벡터가 가장 검출 율이 높았으며 10db에서는 MFCC가 82.03%, LPCC가 15db, 20db에서 각각 93.68%, 93.03%로 가장 검출 율이 높았다. 평균적인 검출율은 LPCC가 81.08%로 가장 높았으며, MFCC가 78.01%, LPC가 77.29%로 비교적 낮게 나왔다.

4.2 사이렌 소리에 대한 실외 환경 실험 결과

실외 환경 실험은 실외 환경에서 발생할 수 있는 위험소리와 실외 환경잡음인 변화가, 골목길, 차도의 잡음을 합성하여 상황별 위험소리 검출을 실험하였다. 실외 위험소리로는 응급차량의 사이렌 소리와 차 경적 소리 2가지에 대해서 실험을 진행하였다.

표 7 사이렌소리에 대한 LPC 특징벡터를 이용한 실험 결과

Table 7 Accuracy result for siren using LPC feature vector

	5db	10db	15db	20db
변화가	55	80	95	100
골목길	60	65	70	85
차도	50	75	80	95

표 7은 LPC 특징벡터를 사용하여 실외 환경별 사이렌 소리를 검출한 실험 결과이다. 5db에서는 평균 55%의 검출 율을 보였으며, 10 db에서는 73.34 %, 15db에서는 81.67%, 20db에서는 93.34 %의 검출 율을 보였다.

표 8 사이렌소리에 대한 LPCC 특징벡터를 이용한 실험 결과
Table 8 Accuracy result for siren using LPCC feature vector

	5db	10db	15db	20db
변화가	60	85	90	100
골목길	60	65	75	85
차도	45	80	80	95

표 8은 LPCC 특징벡터를 사용하여 실외 환경별 사이렌 소리를 검출한 실험 결과이다. 5db에서는 평균 55%의 검출 율을 보였으며, 10db에서는 76.67%, 15db에서는 81.67%, 20db에서는 93.3%의 검출 율을 보였다.

표 9 사이렌소리에 대한 MFCC 특징벡터를 이용한 실험 결과
Table 9 Accuracy result for siren using MFCC feature vector

	5db	10db	15db	20db
변화가	15	15	25	30
골목길	10	25	20	25
차도	10	15	30	35

표 9는 MFCC 특징벡터를 사용하여 실외 환경별 사이렌 소리를 검출한 실험 결과이다. 5 db에서는 평균 11.67 %의 검출 율을 보였으며, 10 db에서는 18.33%, 15db에서는 25 %, 20 db에서는 30 %의 검출 율을 보였다.

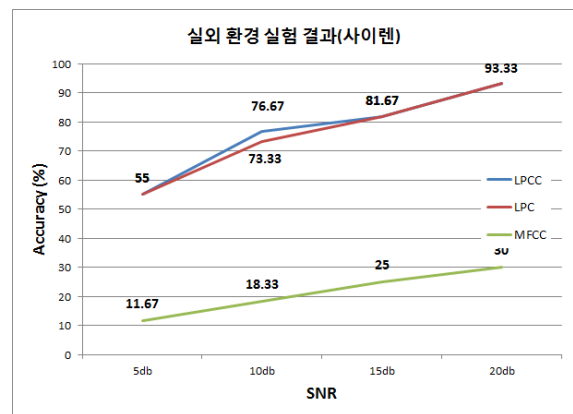


그림 4 사이렌 소리에 대한 실외 환경 실험 결과

Fig. 4 Accuracy comparative experiment result of various SNR for siren

그림 4는 사이렌 소리에 대한 실외 환경 실험에 대한 결과이다. 5db, 15db, 20db에서 LPC, LPCC 특징벡터가 각각 55%, 81.67%, 93.33% 가장 검출 율이 높았으며 10db에서는 LPCC가 76.67%로 가장 검출 율이 높았다. 평균적인 검출 율은 LPCC가 76.67%로 가장 높았으며, LPC가 75.83%, MFCC가 21.25%로 낮게 나왔다.

4.3 차 경적 소리에 대한 실외 환경 실험 결과

표 10 차 경적 소리에 대한 LPC 특징벡터를 이용한 실험 결과
Table 10 Accuracy result for car horn using LPC feature vector

	5db	10db	15db	20db
변화가	60.87	69.57	91.30	91.30
골목길	86.96	86.96	91.30	91.30
차도	73.91	82.61	82.61	86.96

표 10은 LPC 특징벡터를 사용하여 실외 환경별 차 경적 소리를 검출한 실험 결과이다. 5db에서는 평균 73.91%의 검출율을 보였으며, 10db에서는 79.71%, 15db에서는 88.42%, 20db에서는 89.86%의 검출율을 보였다.

표 11 차 경적 소리에 대한 LPCC 특징벡터를 이용한 실험 결과

Table 11 Accuracy result for car horn using LPCC feature vector

	5db	10db	15db	20db
변화가	73.91	91.30	95.65	95.65
골목길	91.30	91.30	95.65	95.65
차도	73.91	86.96	91.30	95.65

표 11은 LPCC 특징벡터를 사용하여 실외 환경별 차 경적 소리를 검출한 실험 결과이다. 5db에서는 평균 55%의 검출율을 보였으며, 10db에서는 10db에서는 89.86%, 15db에서는 94.20%, 20db에서는 95.65%의 검출율을 보였다.

표 12 차 경적 소리에 대한 MFCC 특징벡터를 이용한 실험 결과

Table 12 Accuracy result for car horn using MFCC feature vector

	5db	10db	15db	20db
변화가	26.09	56.52	82.69	100
골목길	78.26	95.65	100	100
차도	60.87	78.26	95.65	100

표 12는 MFCC 특징벡터를 사용하여 실외 환경별 차 경적 소리를 검출한 실험 결과이다. 5db에서는 평균 55.07%의 검출율을 보였으며, 10db에서는 76.81%, 15db에서는 92.75%, 20db에서는 100%의 검출율을 보였다.

그림 5는 차 경적 소리에 대한 실외 환경 실험에 대한 결과이다. LPCC 특징벡터가 5db, 10db, 15db에서 각각 79.71%, 89.86%, 94.2%로 가장 검출율이 높았으며, 20db에서는 MFCC가 100%로 가장 검출율이 높았다. 평균적인 검출율은 LPCC가 89.86%로 가장 높았으며, LPC가 82.97%, MFCC가 81.16%로 비

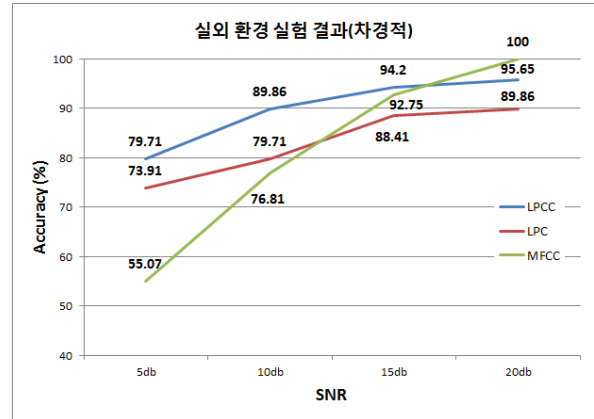


그림 5 차 경적 소리에 대한 실외 환경 실험 결과

Fig. 5 Accuracy comparative experiment result of various SNR for car horn

교적 낮게 나왔다.

3. 결론 및 토의

최근 빠르게 진행되고 있는 고령화 사회로 인해 노인 인구가 늘어나고 난청인구가 증가하고 있다. 청각장애인들은 소리를 듣고 상황을 판단하는 것이 어렵기 때문에 실내, 실외 환경에서의 위험한 상황을 인지 할 수 없어 즉각적인 대처가 불가능하다. 따라서 위험상황에서 빠르게 대처할 수 있는 소리 정보를 청각장애인들에게 전달 해줄 수 있는 보조기기의 연구와 개발이 필요하다. 본 논문에서는 ANN(Artificial Neural Network)을 이용한 청각장애인을 위한 위험상황 검출 엔진을 설계한다. 이를 위하여 위험상황을 실내 위험상황, 실외 위험상황으로 나누고 실내 위험상황에는 물 끓는 소리, 화재경보, 초인종, 전화벨로 정의하였고 실외 위험상황에는 사이렌, 차 경적 소리로 정의하였다. 검출 엔진 실험을 위해 환경잡음을 조용한방, 골목길, 변화가, 차도 총 4가지로 정의하고 약 5분가량의 데이터를 수집하였다. 수집된 위험상황별 데이터를 환경잡음과 합성하였고 신호처리 분야에서 가장 많이 쓰이는 특징벡터인 LPC, LPCC, MFCC를 사용하여 위험상황 검출 정확도를 비교 분석하였다. 그 결과, LPC, LPCC, MFCC 특징벡터간의 분리도 비교 실험에서 LPC 특징벡터의 평균 분리도는 2.791269로 나왔으며, MFCC 특징벡터의 평균분리도는 3.771626으로 나왔다. 가장 높게 나온 LPCC 특징벡터의 평균 분리도는 3.789018로 나왔다. 이는 LPCC 특징벡터가 위험소리 검출에 가장 적합한 특징벡터라고 할 수 있다. 본 논문에서는 실험을 위해 실내 환경에서 발생하는 4가지 위험상황과 조용한 방에서 녹음한 실내 환경잡음과 5db, 10db, 15db, 20db로 합성하여 실험하였고 실외 환경에서 발생하는 2가지 위험상황과 변화가, 골목길, 차도 3가지 환경잡음과 각각 5db, 10db, 15db, 20db로 합성하여 실험하였다. 검출정확도는 LPCC 특징벡터가 5db, 10db, 15db, 20db에서 평균 86.52%로 가장 높았으며, LPC 특징벡터는 81.84%의 검출율을 보였다. 반면, MFCC 특징벡터는 사

이런 소리 검출 율이 평균 21.25%밖에 되지 않아 63.84%의 검출 율을 보였다. 사이렌 소리를 제외한 MFCC의 검출 율은 85.15%로 두 번째로 높은 검출 율을 보였다.

LPC 특징벡터를 추출하는데 사용하는 알고리즘인 Levinson Durbin 알고리즘의 경우 $O(n)$ 의 시간복잡도를 가지고 있으며 [12], LPCC 특징벡터를 추출하는 것 또한 LPC 특징벡터를 기반으로 추출하기 때문에 $O(n)$ 의 시간복잡도를 가지게 된다. 하지만 MFCC 특징벡터를 추출할 때 푸리에 변환이 사용되기 때문에 $O(n^2)$ 이상의 시간복잡도를 갖게 된다. 여기에서 n 은 윈도우 크기이며, 윈도우 크기가 커질수록 MFCC는 시간복잡도가 높아지기 때문에 웨어러블 디바이스와 같은 임베디드 환경에서는 LPC, LPCC 특징벡터가 더 적합하다.

향후, 청각장애인을 위한 위험 검출 엔진을 실제 웨어러블 기기에 적용하여 실제 환경에서의 각 특징벡터의 검출 시간 및 검출 정확도 실험에 대한 연구가 필요하다고 판단된다.

감사의 글

본 연구는 2016년도 상명대학교 교내연구비를 지원받아 수행하였음.

References

[1] Jae-Hun Choi and Joon-Hyuk Chang “Sound Reinforcement Based on Context Awareness for Hearing Impaired” The Institute of Electronics and Information Engineers Vol.48, No.5, pp.109-114, September, 2011.

[2] Y. Toyoda, J. Huang, S. Ding and Y. Liu. “Environmental sound recognition by multilayered neural networks” in Proc. International Conference on Computer and Information Technology, pp. 123-127, September, 2004.

[3] Ji-Eun Kim and Ho-Sub Yoon. “Three sounds classification to recognize dangerous situation for the hearing-impaired” The conference on Journal of the HCI Society of Korea, pp. 570-572, January, 2011.

[4] Oppenheim, Schafer, “discrete-time signal processing 3 edition”, Pearson, 2007.

[5] Haykin, Simon, adaptive filter theory (5th edition), Pearson, 2013.

[6] Y. Tyoda, J. Huang, S. Ding and Y. Liu, "Environmental sound recognition by multilayered neural networks", in Proc. International Conference on Computer and Information Technology, pp.123-127, 2004,

[7] Sung-Woo Byun, So-min Lee, Seok-Pil Lee, “A Selection of Optimal EEG Channel for Emotion Analysis According to Music Listening using Stochastic Variables”, KIEE, Vol.

62, No. 11, pp 1598-1603, 2013.

[8] Ha-Na Choi, Sung-Woo Byun and Seok Pil Lee, “Discriminative Feature Vector Selection for Emotion Classification Based on Speech”, The Transactions of The Korea Institute of Electrical Engineers, Vol.64, No.9, pp.1363-1368 2015

[9] So-Min Lee, Sung-Woo Byun, Seok-Pil Lee, “Comparison of EEG Feature Vector for Emotion Classification according to Music Listening”, KIEE, Vol. 63, No. 5, pp 696-702, 2014.

[10] Kazuhide Okada, Gwan Kim, Pyong Sik Pak, “Sound Information Notification System by Two-Channel Electrotactile Stimulation for Hearing Impaired Persons”, 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp.3826 - 3829 2007.

[11] cs231n, <http://cs231n.github.io/convolutional-networks/>, (accessed June 20, 2016).

[12] Wikipedia, https://en.wikipedia.org/wiki/Levinson_recursion, (accessed June 20, 2016).

저 자 소 개



변 성 우(Sung-Woo Byun)

2014년 상명대학교 디지털미디어학과 이학사
2014년~현재 상명대학교 컴퓨터과학과 석·박사 통합과정
<주관심분야> 멀티미디어처리, 인공지능, 음성신호처리



이 석 필(Seok Pil Lee)

1990년 연세대학교 전기공학과 공학사
1992년 연세대학교 전기공학과 공학석사
1997년 연세대학교 전기공학과 공학박사
1997년~2002년 대우전자 영상연구소 선임연구원
2002년~2012년 KETI 디지털미디어연구센터

센터장

2010년~2011년 미국 Georgia Tech. 방문연구원
2012년~현재 상명대학교 미디어소프트웨어학과 교수
<주관심분야> 멀티미디어 검색, 방송통신시스템, 인공지능