

교통사고음 기반 교통사고검지 알고리즘 개발에 관한 연구

Development of an Automated Traffic Accident Detection Algorithms by processing the traffic sound

유성준

(서울시립대 교통공학과 박사과정)

김영찬

(서울시립대 교통공학과 교수)

권오현

((주)하이테크시스템 연구원)

이건호

(도로교통안전관리공단 책임연구원)

김완기

(동국대 안전공학과 박사과정)

강민석

(서울시립대 교통공학과 석사과정)

목 차

I. 서론

II. 교통사고음 검지 관련 이론

III. 교통사고 검지 알고리즘 개발

1. 교통사고음 검지 방법
2. 교통사고음 특징 추출 데이터 분석
3. GMM(Gaussian Mixture Model)을 이용한 교통사고 검지 알고리즘 개발

IV. 교통사고 검지 알고리즘 검증 평가

1. 검증 평가 방법
2. 검증 평가 결과

V. 결론 및 향후 연구과제

참고문헌

I. 서론

2005년 우리나라의 교통사고는 총 214,171건이 발생하여 342,233명의 부상과 6,376명이 사망하였는데, 이 기준으로 교통사고 처리를 위해 지출된 비용을 추계한 결과 약 9조 1천 2백 29억 원으로 추계되었다.

위와 같이 교통사고로 받는 물질적, 정신적 피해는 매우 심각한 수준으로 세계 각국에서 교통사고방지를 위한 많은 연구들이 개발되고 시행되고 있다.

그 중 교통사고자동기록장치는 교통사고가 잦은 지점에 설치하여 교통사고 발생상황을 객관적으로 기록하는 장치로서 현재 일본을 비롯해 미국, 호주 등지에서 많은 연구가 진행되고 실제 현장에 적용하여 명확한 사고원인분석을 실시하고 있다.

국내에서도 2001년도부터 연구를 시작하여 2004년도에 영상검지를 기반으로 한 교통사고자동기록장치가 서울시 4개 교차로에 시범운영되었고 현재는 정상 운영 중에 있다.

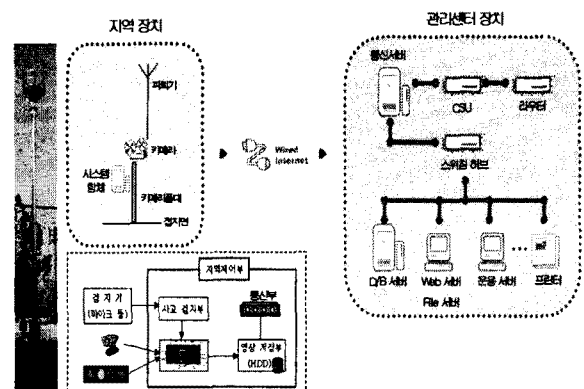
일반적으로 교통사고자동기록장치의 검지기술은 영상검지기술과 음향검지기술을 활용하여 교통사고 발생상황을 자동 검지하는 기술을 접목한 것으로 국가별로 사용되고 있는 검지기술

은 다음과 같다.

<표 1> 국가별 교통사고 검지기술 비교

구분	음향검지기술	영상검지기술
기초이론	FFT, DWT, RCT 등 주파수 필터링 기술	영상 필터링을 포함한 경계추출, 객체 추출 움직임 검지
검지방식	주파수 분리와 DB 필터링	경계추출, 객체추출
적용국가	일본, 영국, 미국	한국(응용기술)

국내에서 사용되고 있는 영상검지기술의 교통사고자동기록장치는 현장에 지역장치와 관리자들이 수시로 검색하고 교통사고 알림서비스를 받을 수 있는 센터로 구성되어 있으며, 개략적인 시스템 구성도는 <그림 1>과 같다.



<그림 1> 현 교통사고기록장치 시스템 구성도

하지만 국내에 적용한 영상검지기법의 교통 사고자동기록장치는 교통사고에 대한 오인식률이 높고 환경변화에 매우 민감하게 작용하는 등의 문제점이 발생하여 영상검지기법의 문제점을 개선하기 위한 방안이 연구되고 있다.

위와 같은 배경에서 본 논문의 목적은 교통 사고음(음향검지) 기반 교통사고 검지 알고리즘을 개발하는데 그 목적이 있다.

본 논문에서는 LPC(Linear Predictive Coding) 계수를 이용한 교통 사고음 특징을 추출한다. 어떤 음향 신호를 인식하기 위해서 그 음향의 특징을 추출하는 것은 가장 기본적이면서도 가장 중요한 부분이다. LPC(Linear Predictive Coding) 알고리즘을 이용하여 해당 음향의 p차원의 특징벡터를 추출한 후, 학습(Training, Modeling)과정을 거쳐 GMM(Gaussian Mixture Model) 특징모델을 생성하고 생성한 GMM (Gaussian Mixture Model) 특징 모델을 기반으로 새로운 음향신호가 입력되었을 때 확률적으로 가장 높은 특징 모델을 얻으므로 교차로 사고 인식을 수행하려고 한다.

II. 교통사고음 검지관련 이론

LPC(Linear Predictive Coding) 계수를 이용한 교통 사고음 특징 추출은 비교적 낮은 차원의 특징벡터로도 시간 영역과 주파수 영역에서의 음향 신호의 특징을 가장 잘 추출해 내는 특징을 가지고 있기 때문이며, 현재 음성 인식 및 합성 분야에 많이 사용되고 있는 알고리즘이기 때문에 본 연구에서 활용하고자 한다.

LPC(Linear Predictive Coding)의 가장 기본적인 개념은 아래의 식에서 언급한 것처럼 시간 n 에서의 주어진 음향 샘플은 과거 p 까지의 음향 샘플의 선형 결합에 의해 근사화 될 수 있다는 것에 근간한다.

$$s(n) \approx a_1 s(n-1) + a_2 s(n-2) + \dots + a_p s(n-p)$$

위의 식에서 a_1, a_2, \dots, a_p 는 음향 프레임 계수이며, 이 계수가 우리가 해당 음향에서 추출하고자 하는 특징벡터이다. 위의 식은 $Gu(n)$ 항을 이용하여 아래와 같이 구체적인

수식화가 용이하다.

$$s(n) = \sum_{i=1}^p a_i s(n-i) + Gu(n)$$

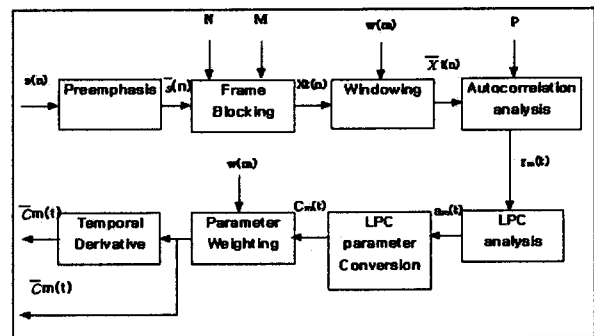
여기서 $u(n)$ 은 단위수열이며, G 는 gain이다. 위의 수식을 z -domain으로 변환하면 아래와 같다.

$$S(z) = \sum_{i=1}^p a_i z^{-i} S(z) + GU(z)$$

$$H(z) = \frac{S(z)}{GU(z)} = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}} = \frac{1}{A(z)}$$

excitation source $u(n)$ 이 G 에 의해 커진 데이터가 all-pole system $H(z)$ 의 입력으로 이용되어 음향 신호 $s(n)$ 이 생성된다.

실제로 LPC(Linear Predictive Coding) Coefficient를 구하는 과정은 <그림 2>와 같다.



<그림 2> LPC Processor Block Diagram

1) Preemphasis

계산의 안정성과 선형 예측을 이용한 분석시 성도의 영향만을 고려하기 위해 사용되며, 고주파 성분을 강조하는 효과가 있다. 보통 fixed first-order system을 널리 사용한다.

$$H(z) = 1 - \overline{a}S(n-1), \quad 0.9 \leq a \leq 1.0$$

이 경우 Preemphasis Network의 Output인 a 는 input인 $s(n)$ 와 다음과 같은 관계에 있다.

$$\overline{s(n)} = s(n) - \overline{a}S(n-1)$$

2) Frame Blocking

Preemphasis된 음향 신호인 $s(n)$ 를 N 개의 샘플로 블록을 나눈다. 인접한 블록은 M 샘플의 차이가 난다. 즉 $N-M$ 개의 샘플이 overlap 된다.

3) Windowing

각각의 프레임의 처음과 끝에서의 불연속을 최소화하기 위해 Hamming Window를 사용한다. $0 \leq n < N$ 의 윈도우 $w(n)$ 를 사용하면 윈도우의 결과로 신호는 다음과 같이 나타낸다.

$$\widehat{xl(n)} = xl(n)w(n), \quad 0 \leq n < N-1$$

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n < N-1$$

각 프레임마다 같은 윈도우를 적용하므로 미리 계산해 테이블로 저장해 둔다.

4) Autocorrelation Analysis

Windowing된 각각의 프레임을 Autocorrelate 한다.

$$r(m) = \sum_{n=0}^{N-1-m} \overline{xl(n)} \overline{xl(n+m)},$$

highest autocorrelation value p 는 LPC Analysis의 차수이다.

일반적으로 p 는 8에서 16까지의 값이고 $p=8$ 을 가장 널리 사용한. Zeroth Autocorrelation은 L th 프레임의 에너지이다. 프레임의 에너지는 음향 검출 시스템의 중요한 파라미터이다.

5) LPC(Linear Predictive Coding) Analysis

각각의 Autocorrelation된 프레임을 LPC (Linear Predictive Coding) parameter set으로 변환하는 단계에서는 LPC(Linear Predictive Coding) 필터 계수를 구한다. 이런 계수들은 성도의 모양에 대한 정보를 주파수 영역을 근간으로 표현한 것이라고 생각하면 된다. 발음이 서로 다른 것은 발음할 때 성도의 모양이

다르기 때문이다. 따라서 성도의 모양에 대한 정보가 발음에 대한 정보라고 볼 수 있다. 또한 몇 차로 분석하느냐에 따라 분석의 성능이 달라진다. 특정 프레임의 LPC(Linear Predictive Coding) 필터 계수는 Durbin의 방법으로 구할 수 있고 다음과 같이 반복적으로 구할 수 있다.

$$E^0 = r(0)$$

$$k_i = \frac{r(i) - \sum_{j=1}^{i-1} a_j^{(i-1)} r(i-j)}{E^{(i-1)}}$$

$$a_i^{(i)} = k_i$$

$$a_j^{(i)} = a_j^{(i-1)} - k_i a_{i-j}^{(i-1)}$$

$$E^{(i)} = (1 - k_i^2) E^{(i-1)}$$

① Durbin의 방법

$$\begin{bmatrix} r_0 & r_1 & \dots & r_{p-1} \\ r_1 & r_0 & & \vdots \\ \vdots & & \ddots & r_1 \\ r_{p-1} & r_{p-2} & \dots & r_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix}$$

$$r_k = \sum_{n=0}^{N-1-k} s(n)s(n+k), \quad 0 \leq k \leq p$$

$a_1, a_2, a_3, \dots, a_p$ 가 LPC계수이다.

② Covariance method

$$E = \sum_{n=0}^{p-1} e_n^2 = \sum_{n=0}^{p-1} (s_n - \sum_{k=1}^n a_k s_{n-k})^2$$

에 대해 E 가 최소가 되는 경우는 기울기가 0이 되는 즉 미분이 0이 될 때이다. 따라서 다음과 같이 쓸 수 있다.

$$\phi_{i,q} = \sum_{k=1}^p \phi_{i,k} a_k$$

이를 Matrix로 나타낸다면,

$$\begin{bmatrix} \phi_{1,0} \\ \phi_{2,0} \\ \phi_{3,0} \\ \dots \\ \phi_{p,0} \end{bmatrix} = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \phi_{1,3} \\ \phi_{2,1} & \phi_{2,2} & \phi_{2,3} \\ \phi_{3,1} & \phi_{3,2} & \phi_{3,3} \\ \dots & \dots & \dots \\ \phi_{p,1} & \phi_{p,2} & \phi_{p,3} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ a_p \end{bmatrix}$$

이와 같이 The Covariance method solution 은 매트릭스 inverse에 의하여 얻을 수 있다.

6) LPC(Linear Predictive Coding) Parameter

Conversion to Cepstral Coefficients

LPC(Linear Predictive Coding) 필터 계수가 구해지면 이를 cepstral 계수로 변환한다. cepstral coefficients란 로그 크기 스펙트럼(Log Magnitude Spectrum)의 FT(Fourier Transform)의 계수를 말한다. 음성인식에서 LPC계수, log area ratio,보다 더 강건하고 믿을 수 있는 것으로 보인다.

$$c_0 = \ln \sigma^2$$

$$c_m = a_m + \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k} \quad 1 \leq m \leq p$$

$$c_m = \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k} \quad m > p$$

일반적으로 $Q > p$ 로 나타난다.

7) Parameter Weighting

cepstral 계수의 민감도를 완화시키기 위해 Tepered Window 기법을 사용한다. 일반적인 방법으로 Log Magnitude Spectrum 과 The Differentiated(In Frequency) Log Magnitude Spectrum을 사용한다.

⊗ Log Magnitude Spectrum :

$$\log |S(e^{j\omega})| = \sum c_m e^{-j\omega m}$$

⊗ Differentiated Log Magnitude :

$$\frac{\sigma}{\sigma \omega} [\log |S(e^{j\omega})|] = \sum (-j\omega) c_m e^{-j\omega m}$$

⊗은 ⊗에서 고정된 스펙트럼의 기울기를 가지고 있다. 따라서 미분을 하면 기울기는 상수가 된다. 또한 ⊕일 때 나타나는 peak값 (e.g. Formants)을 미분을 해도 잘 가지고 있다. 즉, 미분을 함으로써 $(-jm)$ 이 곱하게 되어 Weighting으로써 나타낼 수 있다.

$$\frac{\sigma}{\sigma \omega} [\log |S(e^{j\omega})|] = \sum_{m=-\infty}^{\infty} \hat{c}_m e^{-j\omega m}$$

$$\text{where } \hat{c}_m = w_m c_m \quad \hat{c} = w_m c_m \quad 1 \leq m \leq Q$$

다음과 같은 윈도우를 통하여 무한한 계산을 줄인다.

$$W_m = \left[1 + \frac{Q}{2} \sin\left(\frac{\pi m}{Q}\right)\right], \quad 1 \leq m \leq Q$$

8) Temporal Cepstral Derivative

Local 스펙트럼의 특성을 나타내기에 cepstral 계수는 매우 훌륭하다. 더 향상되고 확장된 스펙트럼의 표현을 위해 Temporal Cepstral Derivative를 사용한다. Log Magnitude Spectrum을 시간에 대해 미분하면 Fourier Series form형태로 나타난다. 즉, 시간 t일 때 m번째의 Cepstral 계수 $C_m(t)$ 는 Discrete Time 으로 표시된다.

$$\frac{\sigma}{\sigma \omega} [\log |S(e^{j\omega}, t)|] = \sum_{m=-\infty}^{\infty} \frac{\sigma C_m(t)}{\sigma t} e^{-j\omega m}$$

$$\frac{\sigma C_m(t)}{\sigma t} = \Delta C_m(t) = \mu \sum_{k=-K}^K k c_m(t+k)$$

μ 는 적당한 Normalization 상수, 이상의 계산에 의하여 각각의 프레임에서 LPC(Linear Predictive Coding)의 계수는 Weighted Cepstral 계수는 Q의 벡터이다. 따라서 Q의 벡터를 첨부해야 한다.

$$Q' = (\epsilon_1(t), \epsilon_2(t), \dots, \epsilon_Q(t), \Delta c_1(t), \Delta c_2(t), \dots, \Delta c_Q(t))$$

t는 2Q의 요소를 가진 벡터이다. Q'는 Transpose를 의미한다.

결과적으로 한 프레임으로부터 다음과 같은 특징 파라미터를 추출해내게 된다.

$$Q \text{개의 가중 Cepstral 계수 : } \hat{c}_m, \quad 1 \leq m \leq Q$$

즉, 한 프레임의 음향 신호로부터 Q개의 성분을 갖는 특징 벡터를 추출해 이것으로 그 프레임의 음향 신호를 대표하는 것이다.

III. 교통사고 검지 알고리즘 개발

앞서 살펴 본 LPC(Linear Predictive Coding) 알고리즘에 의해 특징벡터를 추출하고, 추출한 특징벡터를 기본으로 각각 분류한 디지털 음향 데이터를 이용하여 학습시켜 GMM 특징 모델을 추출하는 방법으로 교통사고 발생 음을 검지하는 알고리즘의 개발 순서는 다음과 같다.

1. 교통사고음 검지 방법

1) 1 STEP

음향 처리의 순서를 살펴보면 먼저 마이크에 의해 교차로 주변의 잡음과 함께 사고 음이 입력되게 된다.

2) 2, 3 STEP

입력된 아날로그 음향은 A/D변환기에 의해 디지털 음향 신호로 변환되는데, 이때 음질(8bit mono signed)과 Sampling rate(22Khz)를 정하여 이 Format에 맞는 디지털 음향 신호로 생성시킨다.

3) 4 STEP

취득한 디지털 음향 데이터를 몇 가지로 분류로 일정 시간동안 수집한 뒤 LAC알고리즘에 의해 특징벡터를 추출하고, 추출한 특징벡터를 기본으로 각각 분류한 디지털 음향 데이터를 이용하여 학습시켜 GMM특징모델을 추출한다.

4) 5 STEP

실시간으로 입력되는 디지털 음향 데이터에 대해 이미 학습과정을 거쳐 생성한 GMM특징모델과 입력되는 특징벡터간의 확률적인 비교에 의해 사고 음향을 분류해 낸다.

5) 6 STEP

사고가 인식되면 영상 저장을 위한 DSP보드에 신호를 보내게 되어 5초 후 영상입력이 종료되고, 영상 데이터가 저장된다.

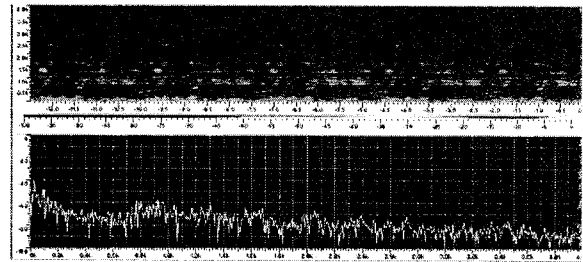
2. 교통사고음 특징 추출 데이터의 분석

교통사고음 수집은 현재 운영중인 16대의 교통사고자동기록장치에 초소형마이크를 설치하여 교통사고음, 클락션, 사이렌, 교통소음 등을 수집을 하였다.

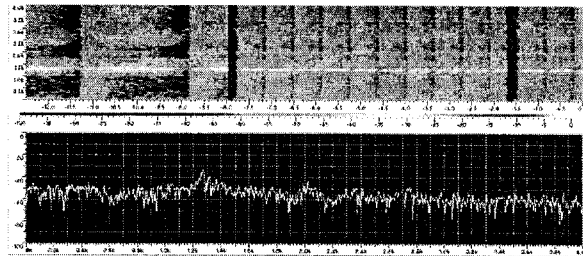
특징이 있는 3가지 종류의 음향 데이터(사고, 정상, 경적 음향 데이터)를 통해 확인해 본 결과 교차로 내의 사고가 발생한 경우에 대해서는 아래와 같이 주파수 영역에서의 주파수의 세기가 전반적으로 커지거나, 혹은 고주파수 대역에서의 주파수 세기가 급격하게 커진다는 것을 확인할 수 있었다. 이는 주파수 영역에서의 신호의 세기를 나타내는 Power Spectrum Density의 임의의 경계 값으로 어느 정도 사고가 발생한 경우를 검지할 수 있으리라 판단되

지만, 본 논문에서는 사고의 종류까지도 인식하려는 목적에서 LPC 계수를 이용하여 특징 벡터를 추출하고, 추출한 특징 벡터와 학습과정을 통해 생성된 특징 모델간의 확률적인 접근에 의해 교차로 내의 검지된 음향 데이터로 사고의 종류를 인식할 수 있으리라 판단된다.

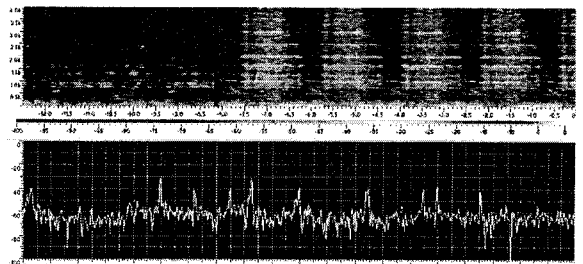
3가지의 음향에 대해 주파수 영역의 spectrum과 spectrogram을 분석해 본 결과, 정상 음향<그림 3> 데이터의 경우에는 사고 음향<그림 4> 데이터에 비해 주파수의 전반적인 세기가 상대적으로 약하다는 것을 알 수 있었고, 경적 음향<그림 5> 데이터의 경우에는 특정한 주파수에서의 세기가 국부적으로 커진다는 것을 확인할 수 있었다(반면 사고 음향 데이터의 경우에는 전반적으로 정상 음향 데이터에 비해 주파수의 세기가 커지고, 주파수 영역에서의 frequency envelope이 매끄럽다는 것을 확인할 수 있다). 그러므로 이러한 주파수적인 특징을 특징벡터를 사용하여 추출하고, 추출한 특징벡터를 통해 교차로 내의 여러 종류의 사고를 인식하고자 한다.



<그림 3> 정상 데이터의 Spectrogram 및 Spectrum



<그림 4> 사고 음향 데이터의 Spectrogram 및 Spectrum



<그림 5> 경적 소리 데이터의 Spectrogram 및 Spectrum

3. GMM(Gaussian Mixture Model)을 이 용한 교통사고 검지 알고리즘 개발

우리는 교차로 부근에서 음향 센서를 통해 추출한 음향 데이터를 몇 가지 종류의 model로 분류시킨다. 몇 가지 종류의 model로 분류한 후 각 데이터들을 30msec단위로 frame을 씌워서 위에서 언급한 LPC(Linear Predictive Coding)를 이용하여 p차원의 특징벡터를 추출한다.

그 결과, 같은 종류의 Model들은 p차원의 공간상에서 서로 유사한 위치에 존재할 것이며, 이 특징들을 우리는 여러 개의 Gaussian pdf Model로 이루어진 GMM(Gaussian Mixture Model)로 학습하여 다음의 특징모델을 추출한다.

$$\mu_{jk}(n) = \frac{1}{N_{jk}(n)} \sum_{i=1}^{N_k(n)} s_{ijk}(n)$$

$$\sigma_{jk}^2(n) = \frac{1}{N_{jk}(n)} \sum_{i=1}^{N_k(n)} (s_{ijk}(n) - \mu_{jk}(n))^2$$

$$\omega_{jk}(n) = \frac{N_{jk}(n)}{N_j(n)}$$

$$, k = 1, 2, \dots, L \quad , j = 1, 2, \dots, M$$

L : n 번째 시간의 j 번째 model의 gaussian pdf 수.

M : 전체 특징모델 수.

$N_{jk}(n)$: n 번째 시간의 j 번째 model의 k 번째 gaussian pdf에 속한 특징벡터 수.

$N_j(n)$: n 번째 시간의 j 번째 model에 속한 전체 특징벡터 수.

$s_{ijk}(n)$: n 번째 시간의 j 번째 model의 k 번째 gaussian pdf에 속한 i 번째 특징벡터.

$\mu_{jk}(n)$: n 번째 시간의 j 번째 model의 k 번째 gaussian pdf의 means.

$\sigma_{jk}^2(n)$: n 번째 시간의 j 번째 model의 k 번째 gaussian pdf의 variance.

$\omega_{jk}(n)$: n 번째 시간의 j 번째 model의 k 번째 gaussian pdf의 가중치.

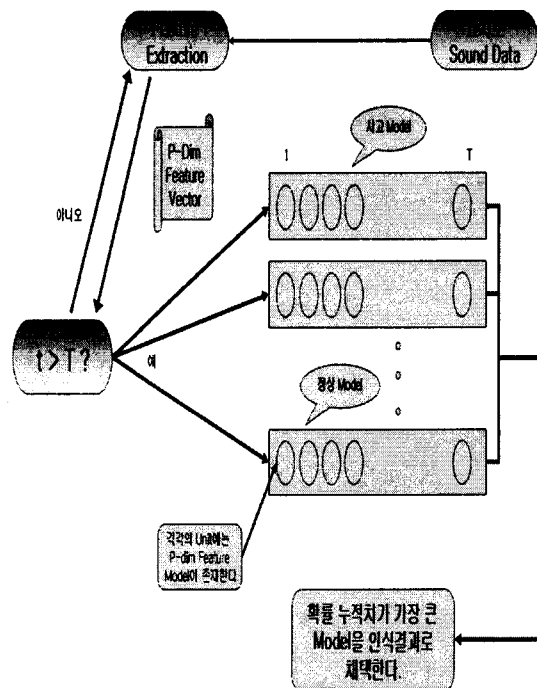
같은 종류의 음향 신호를 다차원(p차원)의

공간상에 투영시킨다면, 다차원 공간상에 상호 유사한 근거리에 존재할 거라는 가정 하에 Gaussian pdf로 이 데이터들을 몇 개의 Gaussian pdf로 Model화 한다는 것은 해당 음향신호만이 보유하고 있는 고유의 어떤 특징들을 Model화 한다는 의미이며, 이러한 Model을 통해 어떤 새로운 특징벡터가 입력되었을 때 확률적인 개념에 의해 원하는 답을 얻는 방법은 여러 인식 분야에서 사용되고 있으며, 우수성이 이미 입증된 방법이기 때문에 우리는 이 논문에서 GMM(Gaussian Mixture Model)을 사용하게 되었다. 위의 각각의 Model에 해당되는 특징모델($\omega_{jk}(n)$, $\mu_{jk}(n)$, $\sigma_{jk}^2(n)$)을 추출한 후, 우리는 다음의 Gaussian pdf로 현재 입력된 음향 데이터를 인식할 수 있다.

$$P_j(s(n)) = \sum_{k=1}^L \frac{\omega_{jk}(n)}{\sqrt{2\pi\sigma_{jk}^2(n)}} e^{-\frac{(s(n) - \mu_{jk}(n))^2}{2\sigma_{jk}^2(n)}}$$

$$J = \operatorname{argmax}_j \sum_{n=0}^T P_j(s(n))$$

위의 식에서 우리는 각 시간에서의 확률 값을 계산했으며, 시간 T 까지의 확률 누적치가 가장 높은 model J 가 우리가 찾은 답이다. 교차로 사고 음향 인식기는 <그림 6>과 같다.



<그림 6> 교통사고음 검지 알고리즘

학습에 사용된 특징벡터는 LPC(Linear Predictive Coding)계수 12차원, LPC(Linear Predictive Coding)에너지 1차원, LPC(Linear Predictive Coding)캡스트럴 계수 12차원, LPC(Linear Predictive Coding)캡스트럴 에너지 1차원을 모두 포함한 26차원의 벡터를 사용하였으며, 학습 데이터는 각각의 음향 종류별로 약 1000개의 세그먼트이션 된 음향 데이터 과일을 이용하여 학습시켰다.

IV. 교통사고 검지 알고리즘 검증 평가

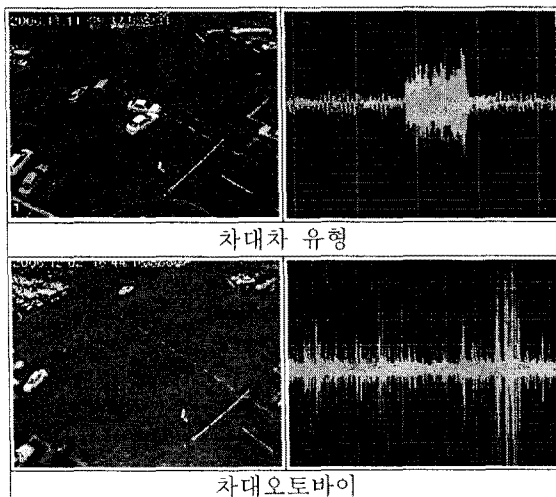
1. 검증 평가 방법

교통사고음 식별 알고리즘 개발에 따른 평가는 현재 서울시와 부산시에 설치된 교통사고자동기록 장치로부터 취득된 교통사고발생음을 개발한 알고리즘에 적용하여 교통사고음 식별율을 자체 실험으로 평가하였다.

취득된 영상을 정상상태, 경적소리, 사고발생으로 크게 3가지로 구분하였으며, 사고발생 상황에는 차대차 사고와 차대 오토바이 사고로 구분하여 각각의 음원에 대한 식별률을 평가하며, 각 상황별 실험데이터 수와 취득된 음원종류는 <표 2>와 같다.

<표 2> 평가를 위한 데이터 수

음원 종류	데이터 수	실험 횟수	
정상 상태	50건	총 2회	
경적 소리	50건	총 2회	
사고			
차대차	25건	총 2회	
발생	차대오토바이	25건	총 2회



<그림 7> 취득된 사고유형별 음원형태

2. 검증 평가 결과

<표 3> 음원종류별 식별 결과

음원종류	총실험횟수	식별률(%)	
정상상태	100회	88%	
경적소리	100회	88%	
사고			
차대차	50회	86%	
발생	차대오토바이	50회	84%

앞에서 언급한 3가지 상태에 대한 실험을 각각 100회 실험한 결과 <표 3>과 같이 식별률이 대부분 최저 84%, 최고 88%로 식별됨을 확인할 수 있었다. 이는 위의 <그림 7>에서 확인할 수 있는 것과 같이 각 음향의 주파수적인 명백한 특징이 존재하므로 특징을 추출하고, 추출한 특징을 모델화한다면 분명 식별할 수 있을 것이라는 가능성에 기인한다. 실제 우리의 실험에서 최소 84%, 최고 88%의 식별 능력을 보인 것은 정상상태, 경적소리, 사고발생음 각각 1000개 음향 학습 데이터를 입력으로 학습시켜 추출한 각각의 GMM(Gaussian Mixture Model)특징 모델들이 위의 3가지 상태에 대한 독립적인 특징을 잘 함축하고 있음을 의미하며, p차원의 공간상에 위의 3가지 학습 데이터들이 대부분 서로 떨어져서 분포하고 있음을 간접적으로 알 수 있다. 위의 식별 결과에서 오식별이 되는 나머지 경우는 3가지 상태에 대해 일부 공통적인 유사성이 있는 특징들 때문에 발생된 것으로 판단되며, 그에 대해서는 3가지 상태에 대해 각각 독립적인 특징이 있는 학습 데이터를 더 많이 이용하여 최적화된 GMM(Gaussian Mixture Model)을 추출하는 것이 중요하리라 판단된다.

V. 결론 및 향후 연구과제

우리는 LPC(Linear Predictive Coding)계수와 GMM(Gaussian Mixture Model)특징 모델을 이용하여 교차로 구간에 설치된 초소형 마이크로 실시간 입력되는 3가지 상태의 음향에 대해 분석 및 식별을 가능하게 한 연구를 수행하였으며, 위의 3가지 상태의 음향 신호를 입력으로 최소 84%, 최고 88%의 식별 결과를 얻을 수 있었다. 이는 비록 단순한 종류의 음

향 신호이지만 음향 신호로도 어느 정도의 교차로 구간에서의 사고 검지를 수행할 수 있음을 의미하며, 향후 교차로 부근에서 발생 가능한 여러 환경에서 획득되는 다양한 음향에 대한 독립적인 특징 모델을 효과적으로 추출하고, 추출한 특징 모델을 기반을 두어 다양성과 범용성 있는 접근을 시도해야 하며, 더불어 교통사고검지와 유효이벤트검지율을 향상시키기 위한 방법의 일환으로 음향검지와 영상검지를 조합 한 사고검지방법을 연구함으로써 각각이 가지고 있는 단점과 한계점을 서로 극복해 줌으로써 교통사고 자동검지율의 향상을 기대할 수 있다.

Detector by Processing the Traffic Sounds, TRB annual Meeting, 2000.

참고 문헌

- [1] 유성준, 김영찬, 이동영, "교차로 교통사고 검지시스템 성능개선 방안 연구", 제5회 한국ITS학회 추계학술대회, 2006
- [2] 도로교통안전관리공단, "교통사고 자동기록장치 효율적 운영방안 연구, 2005. 12.
- [3] 일본 경찰청, "교통사고 자동기록장치의 정비 사업평가결과 보고서", 2003. 12.
- [4] 한민홍 외, "교차로 사고 및 차량추돌시 획득을 위한 영상블랙 박스", 2000년도 지능형교통시스템(ITS) 연구·개발사업 최종보고서
- [5] C. Harlow, Y. Wang, "Automated Accident Detection System", In Transportation Research Record. No. 1746, PP. 90-93, 2001.
- [6] L. M. Bruce, N. Balraj, Y. Zhang, Q. Yu, "Automated Accident Detection in Intersections via Digital Signal Processing". Presented at 83rd Annual Meeting of the Transportation Research Board, 2004.
- [7] L.R.Rabiner, "Fundamentals of speech recognitions", Prentice Hall, 1994.
- [8] Naoaki Suetsugi, Hideki Nakanmichi, Masahiro Toya, "New Traffic Incident