

생체기반 GMM Supervector Kernel을 이용한 운전자검증 기술

Driver Verification System Using Biometrical GMM Supervector Kernel

김형국*
(Hyoung-Gook Kim)

요약

본 논문에서는 음성과 얼굴 정보를 분석하여 자동차환경에서 운전자를 검증하는 기술을 소개한다. 음성정보를 이용한 화자검증을 위해서는 잘 알려진 Mel-scale Frequency Cepstral Coefficients(MFCCs)를 음성 특징으로 사용하였으며, 동영상 을 이용한 얼굴검증에 대해서는 AdaBoost를 이용하여 검출된 얼굴 영역에 대해 주성분 분석을 수행하여 데이터의 크기가 현저히 줄어든 특징벡터를 추출하였다. 기존의 화자검증 방식에 비해 본 논문에서는 추출된 음성 및 얼굴 특징들을 Gaussian Mixture Models(GMM)-Supervector기반의 Support Vector Machine(SVM)커널 방식에 적용하여 운전자의 음성과 얼굴을 효과적으로 검증하는 방식을 제안하였다. 실험결과 제안한 방법은 단순한 GMM 방식이나 SVM 방식보다 운전자 검증성능을 향상시킬 수 있었다.

Abstract

This paper presents biometrical driver verification system in car experiment through analysis of speech, and face information. We have used Mel-scale Frequency Cepstral Coefficients (MFCCs) for speaker verification using speech information. For face verification, face region is detected by AdaBoost algorithm and dimension-reduced feature vector is extracted by using principal component analysis only from face region. In this paper, we apply the extracted speech- and face feature vectors to an SVM kernel with Gaussian Mixture Models(GMM) supervector. The experimental results of the proposed approach show a clear improvement compared to a simple GMM or SVM approach.

Key words: Driver verification, adaboost algorithm, SVM kernel, gaussian mixture models

I. 서론

지문이나 손 모양, 음성, 홍채, 망막, 혈관, 서명에 이르기까지 개인에 따라 그 특징이 명확하게 다른 신체 부위나 행동 특성으로 사람을 식별하고 검증 하는 생체인식 기술은 전통적인 보안시스템을 대신할 차

세대 보안 기술로 각광받고 있다. 최근 들어 이러한 기술을 기반으로 등록된 운전자를 인식 및 검증하여 등록된 운전자만이 시동 후 운전이 가능하게 하는 차량도난 방지 시스템에 대한 연구들이 진행되고 있다.

일반적으로 얼굴 인식 및 검증 시스템은 먼저 주어진 이미지 내에서 추출된 색상과 에지 성분을 이

* 주저자 및 교신저자 : 광운대학교 전자융합공학과 부교수

† 논문접수일 : 2009년 12월 30일

† 논문심사일 : 2010년 4월 20일

† 게재확정일 : 2010년 4월 25일

용하여 얼굴영역을 검출한다. 다음 단계로 검출된 얼굴영역에 대한 주성분 분석을 수행하여 데이터의 크기를 현저하게 줄인 특징벡터를 GMM(Gaussian Mixture Model)이나 SVM(Support Vector Machine) 등의 분류방식에 사용하여 누구인지를 판단하는 방식들이[1] 사용되어 왔다. 그리고 음성정보를 이용한 화자인식 및 검증에 있어서는 입력된 음성신호로부터 추출된 MFCC(Mel-scale Frequency Cepstral Coefficients) 특징벡터를 SVM에 적용하는 방식들이[2] 제안되어 왔다.

본 논문에서는 GMM 방식과 SVM 방식을 효과적으로 결합한 GMM supervector kernel 방식을 통해 음성검증과 얼굴검증을 수행한다. 그리고 각 검증결과를 결합하여 자동차 환경에서 운전자를 인식하는 검증항상 알고리즘을 제안한다.

본 논문의 구성은 다음과 같다. II 장에서는 제안된 운전자검증 시스템에 대해서 소개하며 III 장에서는 제안된 방식을 적용한 실험결과를 소개하고 마지막으로 IV 장에서 결론을 맺는다.

II. 운전자인증 시스템의 구성

본 논문에서 사용한 운전자인증 시스템은 <그림 1>에 나타나 있다.

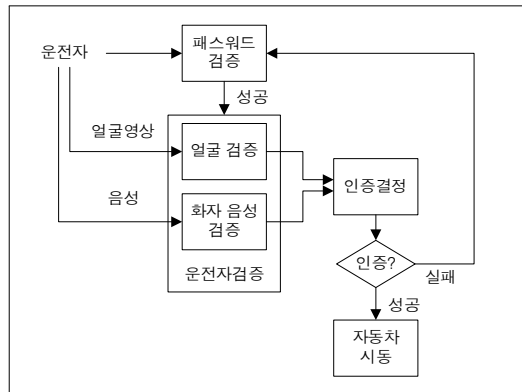
본 논문에서 제안된 생체기반 운전자검증 시스템은 일반적인 검증시스템과 마찬가지로 패스워드 검증 구간이 기본적으로 사용되며, 추가적으로 생체기반의 사용자 정보로서 자동차 내의 카메라와 마이크로로부터 입력받는 얼굴 영상과 음성 데이터의 특징을 사용한다. 즉, 사용자가 운전좌석에 앉으면 시스템은 패스워드를 요구하고, 패스워드가 일치하면 얼굴 검증 단계와 음성 검증 단계로 들어간다. 최종적으로 얼굴 검증 결과와 음성 검증 결과를 결합하여 설정된 운전자 검증 임계값과 비교해서 운전자의 검증을 수락할 것인지, 거절할 것인지를 결정한다.

1. 얼굴검증 알고리즘

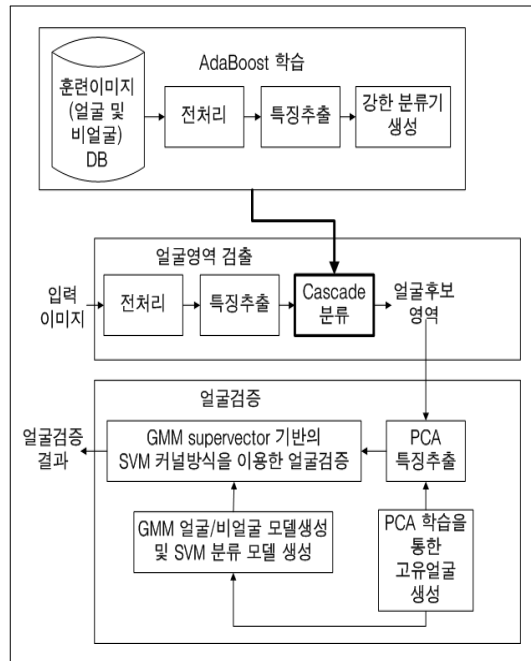
본 논문에서 사용한 얼굴인식 알고리즘은 <그림

2>에 나타난 바와 같이 크게 학습단계, 얼굴영역 검출단계와 얼굴검출단계로 나뉜다.

AdaBoost 학습단계에서는 얼굴과 비얼굴 패턴으로 구성된 훈련영상 집합에서 전처리 단계로 명암대비 스트레칭 기법과 이미지 내의 명암 분포를 보정하는 히스토그램 평활화 기법을 통해 명암 대비도를



<그림 1> 생체기반 운전자검증 시스템 구조
<Fig. 1> Block diagram of a biometrical driver verification.



<그림 2> 얼굴검증 알고리즘 구조
<Fig. 2> Block diagram of a face verification algorithm.

키움으로서 얼굴과 배경의 경계를 더욱 부각시킨다. 동영상 내의 얼굴의 움직임이 발생한 경우에 이전 프레임과 현재 프레임의 차이를 이용하는 모션 추적 과정 등의 전처리 단계를 수행한 후에, 얼굴영역을 검출하기 위해 간단하면서도 연산이 빠른 Haar-like 특징[3]을 추출한다. Haar-like 특징은 인테그랄 이미지를 이용하여 빠르게 연산할 수 있을 뿐만 아니라, 유동적인 얼굴의 특성에도 높은 검출 성능이 요구되는 실시간 객체 검출에 효과적으로 적용된다. Ada-Boost기반의 학습 알고리즘은 오분류된 훈련 패턴의 가중치를 증가시키고, 빠르게 분류된 훈련 패턴의 확률분포를 감소시키는 과정을 반복한다. 이 과정을 통해 각 단계마다 특징의 선형집합 중에서 얼굴을 검출하는데 결정적인 역할을 하는 특징만을 남기고 나머지는 제거함으로써 최소의 에러율을 나타내는 Haar-like 특징만을 선택한다.

개의 약한 분류기의 선형적인 결합을 통해 최종적으로 높은 검출성능을 가지는 강한 분류기를 생성하는 AdaBoosting 기법과 cascade 구조 [3]에 생성된 얼굴 특징 계수를 적용하여 입력된 동영상으로부터 얼굴후보 영역을 검출한다.

검출된 얼굴 후보 영역으로부터는 주성분 분석인 Principal Component Analysis(PCA)를 통해 전체 영상의 데이터를 데이터의 분산이 큰 몇 개의 고유얼굴에 대한 축으로 선형 투사시켜 데이터의 타원을 줄임으로써, 얼굴영역 이미지의 특징 파라미터를 추출한다. 고유 얼굴 생성은 얼굴영역집합으로부터 PCA 학습을 수행하여 획득된다.

검출된 얼굴영역으로부터 PCA를 통해 추출된 특징들은 얼굴검증을 위해 GMM supervector[4]를 SVM Kernel에 적용한 GMM supervector kernel을 사용하여 학습과 검증을 수행한다. 상세한 학습 및 검증과정은 GMM supervector kernel를 이용한 SVM에서 설명된다. SVM은 이진 분류기로서 2개의 클래스를 분류하기 위한 방법으로 한 클래스와 나머지 다른 클래스들을 학습하는 방법이다. 각각의 이진 SVM의 결과로 가장 높은 값을 갖는 경우 해당 클래스로 분류하게 된다.

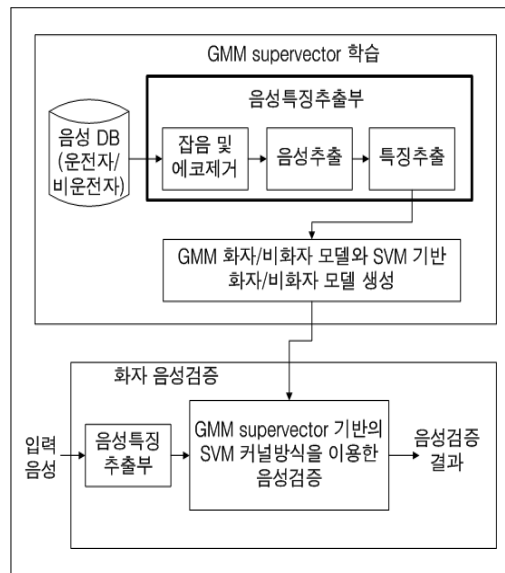
2. 화자인증 알고리즘

화자인식 알고리즘은 크게 GMM supervector 학습부와 화자음성검증부로 구성되며 전체적인 흐름도는 <그림 3>과 같다.

GMM supervector 학습부에서는 운전자와 그 외의 화자들에 대한 음성집합들로부터 잡음과 에코성분에 강인한 음성특징을 추출한다. 그리고 추출된 음성특징들로부터 GMM기반의 화자 및 비화자 모델을 생성한다. 생성된 화자모델간의 kernel을 SVM기반의 학습을 통해 화자 및 비화자 경계구간을 결정하는 support vector를 검출한다.

본 논문에서는 [5]에서 사용한 음성향상 알고리즘을 사용하여 잡음과 에코를 제거하는 방식을 사용한다. 즉, 첫 단계로 크로스 스펙트럼 추정에 기반한 적응필터를 통해 에코를 제거하고, 두 번째 단계로 Generalized Gamma 분포기반의 Log Spectral Amplitude 음성추정 방식 추정을 통해 외부 배경잡음을 제거하여 음성의 음질을 향상시킨다.

다음 단계로 이중 버퍼를 이용하여 실시간으로 잡음과 음성이 제거된 음성 데이터를 입력 받으면서



<그림 3> 음성검증 알고리즘 구조
 <Fig. 3> Block diagram of a speaker verification algorithm.

절대 에너지 방식을 이용하여 음성인지 잠음인지를 체크하는 단계를 수행한다. 즉, 각 프레임의 절대 에너지가 설정된 값보다 크면 이 구간을 실제 음성으로 간주한다. 본 논문에서는 음성 프레임의 길이를 100으로 설정하고, 데이터의 손실을 방지하기 위해서 프레임을 중복시키는 방식인 프레임 블록킹을 30으로 적용하였으며, 음성구간 추출은 절대 에너지가 설정된 값보다 큰 프레임이 6개 이상이 되면 그 부분은 유성을 구간으로 간주하여 메모리에 저장한다. 반면에, 절대 에너지가 설정된 값보다 작은 프레임이 20개 이상이 되면 무성음 구간까지 음성의 끝점을 검출하여 저장한다. 아울러 끝점이 검출되면 20 프레임의 무성음 구간 중 15 프레임은 버리고, 나머지 5 프레임만을 음성영역으로 사용한다[6].

검출된 음성프레임은 특징 추출 알고리즘인 MFCC 방식을 통해 음성의 특징 정보를 추출한다. 이 절차는 pre-emphasis 처리, 해밍 윈도우 처리, fast Fourier transform을 각각 수행하여 주파수영역으로 변환한다. 변환된 값들은 MFCC 대역 필터뱅크를 통과시키고, 로그화, discrete cosine transform을 거치면 프레임당 13개의 계수값을 획득하는데, 이 값들이 음성의 특징 정보들이며 식 (1)과 같이 표현된다.

$$c(n) = \sum_{k=1}^M \log(S(k)) \cos \left[(k-0.5) \frac{\pi}{M} \right], \quad (1)$$

$$n = 1, \dots, L$$

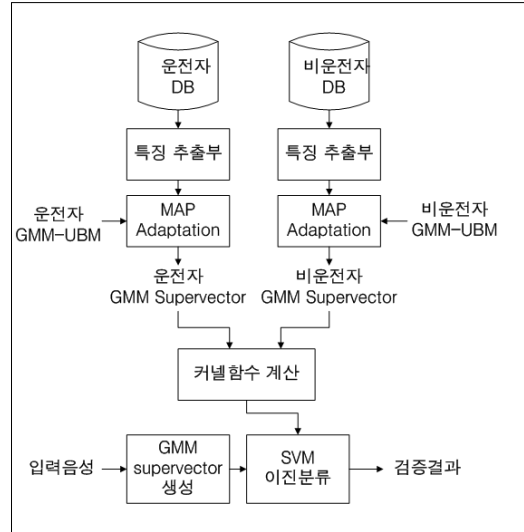
여기서, $S(k)$ 는 k 번째 Mel-filter bank를 통과한 신호이고, M 은 필터뱅크 개수, L 은 MFCC 차수이다.

추출된 음성특징들은 화자 음성 검증을 위해 GMM supervector kernel을 통해 학습과 검증을 수행한다.

3. GMM Supervector Kernel을 이용한 SVM

본 논문에서는 얼굴검증 및 화자검증의 정확도를 향상시키기 위해 GMM 방식과 SVM 방식을 결합한 GMM supervector kernel을 사용하며, 그에 대한 알고리즘 구조 <그림 4>와 같다.

화자검증과 얼굴검증의 방식이 동일하기 때문에 본 소단원에서는 화자검증의 경우를 통해 GMM



<그림 4> GMM supervector kernel을 이용한 SVM 알고리즘 구조

<Fig. 4> Block diagram of a SVM using GMM supervector kernel

supervector kernel을 이용한 SVM 알고리즘을 설명한다. 음성 DB는 운전자 음성 DB와 비운전자 음성 DB로 구성되며, 다양한 잠음 환경 하에서 녹음된 음성정보를 사용하여 GMM 학습을 통해 운전자 GMM-Universal Background Model(GMM-UBM)과 비운전자 GMM-UBM을 각각 생성한다.

GMM-UBM은 여러 개의 가우시안 확률밀도 함수들에 각각의 가중치를 준 다음, 이를 선형 결합함으로써 임의의 모양을 갖는 확률밀도 함수를 표현할 수 있으며 다음과 같은 식으로 표시된다.

$$g(x) = \sum_{i=1}^N \lambda_i \mathcal{N}(x; m_i, \Sigma_i) \quad (2)$$

여기서, λ 는 결합 가중치(mixture weight), $\mathcal{N}(\cdot)$ 는 F 개의 가우시안 분포이고, m 와 Σ 는 가우시안 분포의 평균 벡터와 공분산 행렬을 각각 나타낸다.

GMM-UBM 운전자 모델은 주어진 운전자 학습 음성 데이터에 대해 MAP adaptation을 수행하여 운전자 GMM-UBM의 유사도를 최대화하는 모델 파라미터인 운전자 GMM supervector를 형성한다. 동일한 과정을 통해 비운전자 GMM supervector가 형성된다.

이렇게 형성된 운전자 GMM supervector와 비운전자 GMM supervector는 SVM 학습에 적용되어 운전자와 비운전자를 식별하는 support vector를 생성한다. 생성된 support vector는 입력된 음성에 대해 화자 음성 검증을 결정하는 SVM기반 검증방식에 적용된다.

SVM은 특정 공간에서 각 범주의 대표적인 특성을 갖는 벡터를 이용하여 최적의 하이퍼분리면(hyperplane)을 찾아서 두 범주를 분류하는 기법으로서 다음과 같이 표현된다.

$$f(x) = \sum_{j=1}^J \alpha_j t_j K(x_j, y_j) + d \quad (3)$$

식 (3)은 하이퍼분리면의 일반식으로 a 는 Lagrangian multiplier, d 는 바이어스, $K(,)$ 는 보다 높은 차원의 맵핑을 수행하는 kernel함수로서, 식 (4)와 같은 Kullback-Leibler divergence 선형커널함수를 통해 입력 데이터 공간을 보다 높은 특징 공간으로 맵핑하여 보다 높은 유클리드 공간으로 학습 데이터를 투영한다.

$$K(x_j, y_j) = \sum_{j=1}^J \left(\sqrt{\lambda_j} \Sigma^{-\frac{1}{2}} m_j^x \right)^t \left(\sqrt{\lambda_j} \Sigma^{-\frac{1}{2}} m_j^y \right) \quad (4)$$

이후의 인증과정은 식 (5)의 좌변의 x 에 테스트 벡터를 대입하여 그 값이 양이면 수락하고, 음이면 거절하게 된다.

$$\sum_{j=1}^J \alpha_j t_j K(x_j, y_j) = 0 \quad (5)$$

III. 실험 및 결과고찰

본 논문에서는 운전자 얼굴검증 방식과 음성검증 방식을 독립적으로 적용한 후에, 각각의 검증결과를 결합하여 최종적으로 운전자인지를 결정하는 실험을 수행하였다.

얼굴영역 검출을 위한 AdaBoost 학습을 위해서는 MIT 얼굴 DB를 이용하였다. 학습이미지의 크기는 24×24로 정규화 하였고 24개의 캐스케이트 구조를

통해 최종 전체 특징계수 6050개를 사용하였다. AdaBoost 테스트 단계에서는 카메라의 영상 320×240 크기의 이미지를 입력하여 얼굴 후보 영역을 검출하였고, PCA 학습단계에서는 32×32 크기를 입력하여 1024 개의 고유 얼굴을 생성하였다.

음성데이터는 마이크를 통해 8kHz/8 비트의 PCM 방식으로 녹음하였으며 끝점검출과 전처리 과정을 거친 후, 20 ms의 프레임 크기와 10ms의 프레임 간격을 사용하여 13차 MFCC 계수, 13차 delta-cepstral 계수를 포함하여 총 26차의 음성특징을 추출하였다. 운전자 및 비운전자 화자모델링은 GMM-UBM 방식에서는 가우시안 믹처의 개수를 2048 개로 고정하여 사용하였으며, SVM 배경은 2325 GMM supervector를 추출함으로써 획득하였다. 추출된 GMM supervector와 SVM 배경을 사용하여 SVM 학습을 수행하였다.

얼굴검증을 위해서는 다양한 실내/외 환경과 주차장에서 자동차 운전석에 앉은 20명의 얼굴을 녹화한 얼굴 DB와 20명이 각각 10문장을 발성한 음성 DB를 구성하여 실험에 사용하였다.

표 1은 본 논문에서 제안한 얼굴검증 방식인 GMM supervector kernel을 이용한 SVM (GMM SK) 방식을 기존의 단순한 GMM, SVM 방식과 성능을 비교하였다.

표 1의 결과를 고찰해 보면 제안된 방식의 얼굴검증 알고리즘이 기존의 방식인 GMM, SVM에 비해 보다 향상된 정확률 개선을 보여줌을 알 수 있다.

<표 1> 운전자검증을 위한 기존방식과 제안된 방식의 개선 결과 비교

<Table 1> Comparison between the proposed method and the conventional methods(GMM, SVM) for driver verification

	GMM	SVM	GMM SK
얼굴검증 정확률	85.5%	89.3%	94.7%
음성검증 정확률	88.2%	90.7%	95.6%
얼굴+음성검증 정확률	92.6%	95.4%	98.5%

음성검증에 있어서도 제안된 음성검증 알고리즘 GMM SK 방식이 기존의 방식인 GMM, SVM에 비해 보다 향상된 정확률 개선을 보여주었다. 최종적으로 제안된 GMM SK 방식을 얼굴검증 결과와 음성검증 결과를 결합하여 운전자검증에 적용한 결과 98.5%의 높은 검출 정확률을 보였으며, 이는 GMM 방식보다는 5.9% 개선되었고, SVM 방식보다는 3.1%가 개선되었음을 알 수 있었다.

IV. 결 론

본 논문에서는 차량도난 방지 및 보안을 위한 운전자검증 알고리즘을 제안하였다. 기존의 검증 알고리즘에 적용되어 왔던 단순한 GMM 혹은 SVM 방식을 효과적으로 결합한 GMM supervector kernel을 이용한 SVM 방식을 통해 기존의 방식들보다 향상된 실험결과를 획득하였다. 제안된 방식의 강인성은 실제 자동차 보안환경에서 매우 유용하게 사용될 수 있으리라 생각된다.

참 고 문 헌

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey,"

ACM Computing Surveys(CSUR), vol.35, no.4, pp. 399-458, Dec. 2003.

[2] W. M. Campbell, J. P. Campbell, D. A. Reynolds, E. Singer, and P. A. Torres-Carrasquillo, "Support vector machines for speaker and language recognition," *Computer Speech & Language*, vol.20, no.2-3, pp. 210-229, Apr. 2006.

[3] P. Viola, and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proc. Computer Vision and Pattern Recognition*, vol.1, pp. 511-518, Dec. 2001.

[4] W. M. Campbell, J. P. Campbell, D. A. Reynolds, E. Singer, and P. A. Torres-Carrasquillo, "Support vector machines for speaker and language recognition," *IEEE Signal Processing Letters*, vol.13, no.5, pp. 308-311, 2006.

[5] H. G. Kim, "Implementation of chip and algorithm of a speech enhancement for an automatic speech recognition applied to telematics device," *한국ITS학회논문지*, 제7권, 제5호, pp. 90-96, 2008. 10.

[6] 권순량, "화자 인증이 포함된 실시간 원격 도어락 제어 시스템 개발에 관한 연구," *퍼지 및 지능시스템학회 논문지*, vol.15, no.6, pp. 714-719, 2005. 12.

저자소개



김 형 국 (Kim, Hyoung-Gook)

2007년 3월~현재 : 광운대학교 전파공학과 부교수