Driver Recognition Using Gaussian Mixture Models and Decision Fusion Techniques

Kristin S. Benli, Remzi Düzağaç, and M. Taner Eskil

Department of Computer Engineering,
Işık University, Şile, Istanbul, Turkey
{kristin,eskil}@isikun.edu.tr,
{remzi.duzagac}@isik.edu.tr
http://pi.isikun.edu.tr

Abstract. In this paper we present our research in driver recognition. The goal of this study is to investigate the performance of different classifier fusion techniques in a driver recognition scenario. We are using solely driving behavior signals such as break and accelerator pedal pressure, engine RPM, vehicle speed, steering wheel angle for identifying the driver identities. We modeled each driver using Gaussian Mixture Models, obtained posterior probabilities of identities and combined these scores using different fixed and trainable (adaptive) fusion methods. We observed error rates as low as 0.35% in recognition of 100 drivers using trainable combiners. We conclude that the fusion of multi-modal classifier results is very successful in biometric recognition of a person in a car setting.

Key words: Recognition, Vehicle, Gaussian Mixture Model, Decision Fusion, Biometrics

1 Introduction

Identification of a person in a vehicle setting using solely behavioral signals is a new research area in which few scientific studies were made. The research in this field has many application areas such as authentication, access control, keyless entry, secure communications and automated personalization of vehicle controls.

Specifically person recognition within a vehicle could provide the following benefits [1]:

1. Vehicle safety: Require authorization before or during the driving of the vehicle to make sure that the current driver is an authorized driver. An unauthorized driver can be denied the control of the vehicle.
2. Vehicle personalization: Adjust the vehicle controls according to the driver's physical and behavioral characteristics. Thus, a safer, more comfortable and more efficient driving environment is obtained and the distraction is minimized.
3. Secure mobile transaction opportunities: Mobile banking using biometric authentication may be an example of these opportunities.

For person recognition, we must quantify the characteristic traits of humans and do recognition using a set of robust and discriminative criteria. In Latin, "bios" means "life" and metric means "measure". Therefore, bio-metrics is the study of methods for recognizing people according to one or more real and measurable physical or behavioral traits. Human beings recognize each other using biological characteristics like face, voice or gait. One of the well-known characteristic of the person is the fingerprint which has been used for more than hundred years for person identification. In this study biometrics-based driver identification using solely driving behavior signals is proposed.

Conventionally biometric systems focus on a single trait, such as fingerprints. However the recognition performance of the biometric systems can be improved by using multiple biometric modalities like multiple fingers, or both the face and fingerprint of a person. These systems are known as *multimodal biometric systems* and they are expected to be more reliable due to the presence of multiple pieces of evidence [2].

In this study we use the Nagoya University CIAIR database [3] for driver recognition task. Driving behavior for each driver is modeled using Gaussian Mixture Models (GMMs). In the literature GMM are frequently encountered in text-independent speaker recognition.They were used in the context of modeling the driving signals for the first time by Igarashi[4]. As Igarashi, we used the well-known Expectation Maximization algorithm for training the GMMs.

The goal of this research is to study various decision fusion techniques. We tested the performance of decision fusion using both fixed and trainable methods. Fixed methods have simple fixed rules to combine information from a set of classifiers. We made experiments using five fixed rules; Maximum, Minimum, Median, Mean and Product rules. Trainable methods have free parameters that can be trained on a separate part of the training data. In this study six trainable combiners were used; Fisher, Linear Discriminant, Nearest Mean, Perl, Parzen and K-Nearest Neighbor methods.

This paper reports the results of our study in driver recognition in a database of 100 subjects. In the next section, behavioral signals that are used in the experiments are described. Section 3 revisits the theory of classification in the context of GMMs. Section 4 introduces the decision fusion techniques that are tested in this study. Section 5 details the experiments and Section 6 presents the experimental results. We conclude with our contributions and future research directions.

2 Behavioral Driving Signals

The Center for Integrated Acoustic Information Research (CIAIR) at Nagoya University collected a multi-modal dataset of behavioral signals inside a vehicle [3]. The developers called this system as Data Collection Vehicle (DCV). During the driving process two categories of data were recorded:

- conversations of the driver subjects and

– driving behavior such as the vehicle speed, engine RPM, accelerator/brake-pedal pressure, and steering-wheel motion

In this study we only used driving behavior signals and we show that these signals are sufficient for person identification. A 100 person subset, 50 female and 50 male, of CIAIR database is used in our analysis. For each subject, there are 5 different driving behavior signals available in the CIAIR database. Each channel was sampled at 1.0 kHz and digitized to a signed 16-bit format.

1. Brake pedal pressure (kgf:kilogram force*[1]):
   0 - 50 kgf is mapped to 0 - 5.0V and linearly digitized in the range 0 - 32767
2. Accelerator pedal pressure (kgf:kilogram force*):
   0 - 30 kgf is mapped to 0 - 5.0V and linearly digitized in the range 0 - 32767
3. Engine speed (rpm:revolution per minute):
   0 - 8,000 rpm is mapped to 0 - 5.0V and linearly digitized in the range 0 - 32767
4. Vehicle speed (km/h:kilometer per hour):
   0 - 120 km/h is mapped to 0 - 5.0V and linearly digitized in the range 0 - 32767.
5. Steering angle (degree):
   -1800 degrees to +1800 degrees is mapped to -5.0 to 5.0V and linearly digitized in the range -32768 to 32767.

3 Modeling with Gaussian Mixture Probability Densities

In one dimension the Gaussian probability density function is a bell shaped curve described by two parameters that are mean μ and variance σ^2. In D-dimensional space it is described in a matrix form as

$$N(x; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^{D/2}\left|\Sigma\right|}} exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\} \tag{1}$$

where μ represents the mean vector and Σ represents the covariance matrix.

Gaussian distribution is usually a good approximation for a class model shape in an appropriately selected feature space. It is a mathematically sound function that could be extended to multiple dimensions. In the Gaussian distribution there is an assumption that the class model is a complete model of one basic class. It fails when the actual model, the actual probability density function, is multimodal. For example, assume that we are searching for different face parts from a picture and there are many basic types of eyes, because the people are from different races. Therefore a single Gaussian approximation model would fail to define a wide mixture of all eye types [5].

Gaussian Mixture Model (GMM) is a mixture of several Gaussian distributions. Therefore it can represent the probability distribution in a class that

[1] 1 kilogram force = 9.80665 Newtons

can be further divided into subclasses. Such a mixture of probability density functions can be represented as a weighted sum of Gaussians

$$p(x;\theta) = \sum_{c=1}^{C} \alpha_c N(x;\mu_c,\Sigma_c) \tag{2}$$

where α_c represents the weight of the component c, $0 \prec \alpha_c \prec 1$ for all components, and $\sum_{c=1}^{C} \alpha_c = 1$. The parameter list

$$\theta = \{\alpha_1, \mu_1, \Sigma_1, ..., \alpha_c, \mu_c, \Sigma_c\} \tag{3}$$

describes a particular Gaussian mixture probability density function.

Estimation of the Gaussian mixture parameters for one class can be viewed as an unsupervised learning of those subclasses which contribute to the overall distribution. This is the case where the samples are generated by individual components of the mixture distribution and there is no knowledge about which sample was generated by which component. Clustering attempts to identify the exact components, but Gaussian mixtures can be used for modeling an aggregate distribution of unknown number of classes[5].

4 Decision Fusion

Fusion is combining information from different sources for improving the performance of a system. The most known example of fusion is the use of different sensors for detecting a target. Even if the information is coming from a single sensor, experts may interpret it differently and reach different conclusions. In this case, many experts may be "consulted" to come up with a single decision with highest confidence.

Fusion can be used for many purposes like detection, recognition, identification, tracking, decision making, etc. Information and decision fusion find application areas in defense, robotics, medicine, space and many others. Fusion processes can be categorized into low, intermediate and high level fusion according to the processing stage where the fusion takes place. Low level fusion is also called as *data fusion*. It combines raw data from the several sources for producing new raw data which is expected to be more expressive than the inputs.

Intermediate level fusion is also called as *feature level fusion*. It combines several features like edges, corners, lines, texture parameters, etc. into a feature set. The features may come from several raw data sources like several sensors or come from the same raw data. These features may be obtained from the several feature extraction methods.

High level fusion is also called *decision fusion*. It combines decisions which come from several experts. In other words, if the experts return a confidence (score) instead of a decision, it is a decision fusion problem. Voting methods, statistical methods and fuzzy logic based methods have been used in the literature for decision fusion.

Decision fusion methods can be divided into two main categories; *fixed rules* and *trainable combiners*. Fixed Rules use simple and unchangeable rules for combining different classifiers' results. Trainable combiner rules have free parameters which can be trained on a separate part of the study data. In other words, trainable combiners are classifiers themselves. These combiners do the classification in the score space instead of the feature space.

Let us assume that *S(i,j)* be the score of person i in the modality j for an input test data x. We drop x from the notation for simplicity. The goal is to get a score value *S(i)* for person i by using combination methods [6].

In this sudy, we will use the following fixed and trainable combiner methods for a comparative analysis of their performances.

Fixed Rules:

1. Maximum Rule : $S(i) = max_j\ S(i,j)$
2. Minimum Rule : $S(i) = min_j\ S(i,j)$
3. Mean (Sum) Rule : $S(i) = sum_j\ S(i,j)$
4. Product Rule : $S(i) = prod_j\ S(i,j)$
5. Median Rule : $S(i) = median_j\ S(i,j)$

Trainable Combiners:

1. Nearest Mean Combiner (NMC): Simple linear combiner that accepts nearest class mean as a classifier output.
2. Fisher Combiner (Fisher): Linear classifier that uses the least squares method for matching features and class labels.
3. Linear Discriminant Combiner (LDC): Linear classifier that models every class as a Gauss distribution with the same covariance matrix.
4. Parzen Combiner (Parzen): It uses Parzen density distribution function.
5. K-Nearest Neighbour Combiner (KNN): A method for classifying objects based on the classes of the closest training examples in the feature space.
6. Perl Combiner (Perl): Linear classifier by linear perceptron. A perceptron can be viewed as a binary classifier.

5 Experiments

We analyzed the performance of decision fusion methods on an in-vehicle driver recognition problem. The experiments were done using a 100 person subset of the Nagoya University CIAIR database. This subset consisted of fifty female drivers and fifty male drivers. During the experiments five different "driving behavior signals" were used.

In the first stage noise removal operation was carried out using a low pass filter. This step was followed by decimation procedure which reduces the original sampling rate for a sequence to a lower rate, which is the opposite of interpolation.

The driving signals of each driver were divided into the 20 equal length parts. First 17 parts were used for training, following 2 parts were held-out, and the final part was used for testing. Using a cross validation procedure, 20 different experiments could be carried out on a single person. The held-out data is a part of available training data that is not used during training or testing, but it is used to tune (i.e train) the decision combiner of the recognition system. Held-out data is also called *validation data*.

We used 8 mixture components for GMMs for modeling the driving signals of each person. Also, background GMMs were trained for each modality. In background model, 16 mixture (twice the number of mixtures) components were used [6]. Background GMMs were used for normalization in likelihood ratio testing for biometric recognition.

In the recognition study, the posterior probabilities of driver identities were determined for each channel in the given test data. The identity with the highest probability is chosen as the identity of the test segment. For each channel, the probability of the chosen identity is called a *score* [4]. The similarity between biometrics data is shown with these scores.

One important issue in classifier combination at the score level is to normalize the scores from each modality before fusion. Typical likelihood ranges can be very different among modalities, hence the log likelihood-ratio scores from different modalities cannot be directly added. We normalized the scores using the mean and standard deviation of likelihood scores that were calculated from the held-out validation data. Normalization can be done using a sigmoid function to map the scores to the (0,1) range.

$$S_k^{'} = \frac{1}{1 + exp(-(S_k - \mu)/\sigma)} \tag{4}$$

In equation 4, S_k represents the old log-likelihood-ratio score for the k^{th} modality and $S_k^{'}$ represents the new score. Also, μ and σ represent mean and standard deviation of old scores obtained on the validation set. After normalization these scores were combined by using fixed and trainable fusion methods. *PRTools* [7] software library is used for evaluating the results and combining the classifiers.

6 Results - Driver Recognition

The decision fusion techniques described in the previous paragraphs can be applied to driver recognition where the GMMs trained for each modality suggest a decision. As a result, we have decisions for each modality and these decisions need to be combined in a way that maximizes the performance. This section will present our findings in the driver recognition problem.

Table 1 presents the performance of driver recognition experiments using the individual modalities (i.e. without decision fusion). The best experiment result was obtained with the accelerator pedal presure.

Table 1. Individual performance results for different modalities

Modality	Percent Error
Brake	90.35
Accel	86.15
RPM	98.05
Speed	97.35
Steering	97.90

The highest error rate was obtained with engine speed [2]. The results shown on Table 1 imply that individual driving signals are not appropriate for biometric identification.

This result is compatible with the findings of Wakita et. al. [8]. They present 3 possible reasons for accelerator pedal pressure preserving the most personal property information:

- The accelerator pedal pressure is the direct result of force exerted by the driver
- It is used more frequently than other controls, such as the break pedal
- The acceleration of the vehicle conveys the driving taste of the operator, whereas other signals (e.g. break pedal pressure or steering wheel angle) are more dependent on the driving conditions.

The next stage of our research is driver recognition using the fixed rules and trainable combiners. In this study four different combinations of modalities were chosen and experiments were done using five different fixed rules and six different trainable combiners.

Fixed rules do not show good performances. Their error rates are very high. The lowest error rate, 77.95 percent, was obtained on Combination 1 [3] using the product rule. This error rate is the best we could achieve using fixed rules. Moreover, product rule shows the best performance in all the experiments and has the lowest error rates. Mean rule follows the product rule with the second best lowest error rates.

Overall, the best experiment results were obtained using trainable combiners and specifically the Parzen and K-NN methods. These methods reached error

[2] Abbreviations are used for different channels. These channels are break pedal pressure, accelerator pedal pressure, engine speed, vehicle speed and steering wheel angle respectively.

[3] Note: In the combine list, comma (,) shows the decision fusion where the posterior probabilities of classifiers are combined.
Comb 1 = Brake, Accelerator
Comb 2 = Brake, Accelerator, RPM, Speed, Steering
Comb 3 = Brake, Accelerator, Steering
Comb 4 = Brake, Accelerator, RPM, Speed

Table 2. Percent error rates using fixed rules and trainable combiners

Methods	Comb1	Comb2	Comb3	Comb4
Max	85.95	91.60	89.30	90.75
Min	81.40	90.55	86.50	90.50
Median	78.95	87.70	86.45	86.10
Mean	78.95	85.95	82.30	85.35
Prod	77.95	84.85	81.20	85.05
Fisher	43.45	16.10	29.50	24.55
LDC	35.50	85.55	48.00	78.70
NMC	66.65	54.65	61.40	60.40
Perl	66.00	38.70	55.05	44.25
Parzen	0.55	0.35	0.65	0.80
KNN	0.55	0.35	0.65	0.75

rates as low as 0.35%. Fisher linear combiner showed good performances when the number of input classifiers are increased such as in the second and fourth combination. Linear Discriminant combiner acted oppositely; it showed better performance when the number of input classifiers are decreased. Nearest Mean combiner is a very robust method but generally higher error rates were obtained during the experiments. Specifically in the first and third combination it reached the highest error rates.

7 Conclusion and Future Work

The ultimate goal of our research is to facilitate vehicle-person interaction. In this paper we present our results in driver recognition using solely driver behavior signals and decision fusion. We carried out driver recognition experiments using a 100 person subset of the CIAIR database. This subset consisted of 50 female drivers and 50 male drivers.

In our experiments, the driving signals such as brake pedal pressure, accelerator pedal pressure, engine speed, vehicle speed and steering wheel angle constitute the feature vectors that define the numerical characterization of the biometrics. We used GMMs for approximating the probability distribution of each of these features for each subject in our database. We trained the GMMs using the Expectation Maximization algorithm. For an input pattern we obtain the similarity scores and then combine the decisions of different "GMM experts" for improving the performance of the recognition system. We tested the performance of several fixed and trainable decision fusion techniques.

As expected, trainable combiners showed much higher performance than fixed rules. Among the fixed rules, the product rule showed slightly higher performance. The mean rule consistently rated as the second best classifier after product rule. In general, the performance obtained with fixed rules is slightly higher than that of individual channel performances.

Among the trainable combiners, the Parzen and K-NN consistently rated better than the others. In Parzen and K-NN, we achieved the best performance (0.35%) when we used all driving signals. We obtained next better performance using the brake pedal and accelarator pedal pressures and again the Parzen and K-NN methods. These figures are very promising and they imply that in-car recognition is possible using solely the actions of the driver, specifically the usage of controller pedals. These results also point out a potential of application in a driver verification scenario in which a person who is not confirmed to be the claimed identity will be denied the controls of the vehicle.

We plan to extend our research to the driver verification task. We also plan to extend our experiments to all 812 driver subjects in the CIAIR database. This has proven to be challenging both in terms of the time and memory required for the training of the GMMs. In a parallel study we plan to apply feature reduction/extraction to driving signals to obtain the best meaningful feature set for the characterization of the driver behavior. This will also give us an opportunity to do a comparative study on intermediate (feature) level and high (decision) level fusion.

References

1. Erdogan, H., Ercil, A., Ekenel, H.K., Bilgin, S.Y., Eden, I., Kirisci, M., Abut, H.: Multi-modal person recognition for vehicular applications. N.C. Oza et al. (Eds.): MCS 2005, LNCS 3541, pp. 366 - 375, Monterey CA, (Jun. 2005)
2. Jain, A.K., Ross, A., Prabhakar, S.: An Introduction to Biometric Recognition. IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics, Vol. 14, No. 1, (January 2004)
3. Kawaguchi, N., Matsubara, S., Takeda, K., Itakura, F.: CIAIR In-Car Speech Corpus: Influence of Driving Status. IEICE Transactions on Information and Systems, Vol. E88-D, No.3, (March 2005)
4. Igarashi, K., Miyajima, C., Itou, K., Takeda, K., Abut, H., Itakura, F.: Biometric Identification Using Driving Behavioral Signals. Proceedings IEEE ICME 2004, Taipei, Taiwan, (June 27-30, 2004)
5. Paalanen, P.: Bayesian Classification using Gaussian Mixture Model and EM Estimation:Implementations and Comparisons. Lappeenranta University of Technology, Department of Information Technology, Lappeenranta, Finland, Tech. Rep., (2004)
6. Erdogan, H., Ozyagci, A.N., Eskil, T., Rodoper, M., Ercil, A., Abut, H.: Experiments on Decision Fusion for Driver Recognition. Biennial on DSP for In-Vehicle and Mobile Systems, Taipei, Taiwan, (June 27-30, 2004)
7. Duin, R.P.W., Juszczak, P., de Ridder, D., Paclik, P., Pekalska, E., Tax, D.M.J.: PRTools, a Matlab toolbox for pattern recognition (2004), `http://www.prtools.org`
8. Wakita, T., Ozawa, K., Miyajima, C., Igarashi, K., Itou, K., Takeda, K., Itakura, F.: Driver Identification Using Driving Behavior Signals. IEICE TRANS. INF. & SYST., Vol.E89-D, No.3, (March 2006)