CLASSIFICATION OF VEHICLE OCCUPANTS USING 3D IMAGE SEQUENCES

Pandu R. Rao Devarakota,Bruno Mirbach, Marta Castillo-Franco *Björn Ottersten*

IEE S.A.,
Zone Industrielle, 2b, Route de Trèves
L-2632 Findel, Luxembourg
pandu.devarakota@s3.kth.se

Signals, Sensors and Systems
Royal Institute of Technology (KTH)
SE–100 44 Stockholm, Sweden

ABSTRACT

The deployment of vehicle airbags for maximum protection requires information about the occupant's position, movement, weight, size etc. Specifically it is desirable to discriminate between adults, children, front- or rear faced child seats, objects put on the seat or simply empty seats. 2D images lack depth information about the object and are very sensitive to illumination conditions. Herein, occupant position classification techniques are developed based on low resolution 3D image sequences. The proposed methods are of low complexity and high reliability allowing real time implementation and meeting the rigorous requirements for passenger safety systems. Features are extracted from the 3D image sequences and a *Sequential Forward Search* (SFS) feature subset selection algorithm is employed to reduce the size of the feature set. Two classification techniques are evaluated, the Bayes quadratic classifier and the polynomial classifier. We present the classification results based on a large set of measurements from the low resolution 3D image sequences. The full scale tests have been conducted on a wide range of realistic situations (adults/children/child seats etc.) which may occur in a vehicle.

1. INTRODUCTION

Airbags are designed for frontal impact crashes, the kind of crashes which account for more than half of all passenger vehicle occupant deaths. Airbags are designed to limit head and chest injuries. But, on the other hand there are some instances where occupants are severely injured not because of a frontal crash, but due to the deployment of airbags. According to the American National Highway Traffic Safety Administration (NHTSA), since 1990, 227 deaths have been attributed to airbags deployed in low-speed crashes. The deaths have included 119 children between the ages 1 and 11, and 22 infants. In these cases, the deployment of the airbag should be based on the type and position of the occupant in order to avoid injury. This requires both detection and classification of the occupant. This proves to be a challenging problem due to large variations in the scenes that can be expected and the reliability requirements.

Occupant detection using three dimensional vision has been studied (see for example [1]). However, these methods are based on stereo vision which are sensitive to varying illumination conditions inside the car. Moreover, extra equipment and processing is required to capture 3D information from stereo images. A special 3D imaging sensor has been developed for this application which is free from illumination problems. The results presented here are based on real 3D images obtained by a 3D camera described in [2].

Bayes classifier is one of the main approaches for classification problems. Basically, it utilizes the statistical distributions of the classes in the feature space and an optimized classification criterion. Working under the correct modeling assumptions, the Bayes classifier provides optimal performance [3]. A potential drawback of this approach is that it is difficult to find the posteriori probabilities without any modeling assumptions. One common way to solve this problem is by assuming that the classes are normally distributed. This leads to a simple structure and the moments of the distribution can be estimated from the data. On the other hand, the normal distribution often only provides a rough approximation of the observed data. Thus, applying the normal distribution may result in a suboptimal behavior of the recognition system and the reliability predicted by theory is not achieved in practise [4].

Another well known statistical method for pattern classification is the polynomial classifier based on polynomial regression and a mean-square functional approximation [4]. The advantage of this technique is that it makes no assumption about the underlying statistical model. Any decision rule that consists of a discriminate function that is a linear combination of scalar functions of the pattern vector may be chosen on the basis of a priori knowledge about the classes.

The paper is organized as follows. In the Section 2

This project is funded by IEE S.A. and Luxembourg International Advanced Studies in Information Technology (LIASIT), Luxembourg.

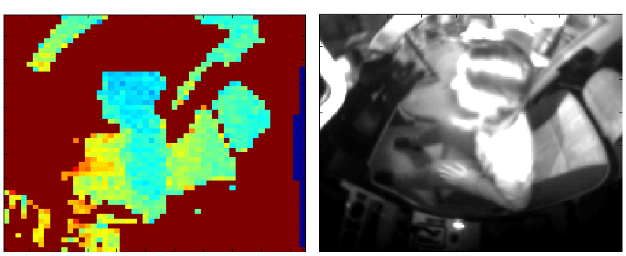
the test sequences used for classification training are described. In Section 3 the feature computation and feature subset selection is presented. In Section 4 the concepts of Bayes quadratic classifier and the polynomial classifier are described. The results in Section 5 provide the main contribution of the paper where the classifiers are compared and validated and Section 6 provides our conclusions and points to future work.

2. THE 3D CAMERA IMAGE DATABASE

To train the classifiers and to perform the recognition task a database of test sequences has been developed at IEE S.A., Luxembourg. At present, the goal is to discriminate between 4 classes, empty, rear faced infant seat (RFIS), forward face child seat (FFCS) and finally an adult person (P). The database consists of 231 sequences of different sizes with a total of 20529 frames of images. In order to take the variation of occupant scenes into account, different occupants with varying hand postures, leg postures, and torso gestures were recorded. For this purpose the camera has been installed in the overhead module of car. The field of view of $120° \times 90°$ allows then a complete survey of the passenger seat. The camera has 50×52 pixels which corresponds to $3\mathrm{cm} \times 2\mathrm{cm}$ lateral resolution (in the average distance of an occupant from the camera). The radial resolution is around 1.0cm. Training data and testing sequences were selected randomly from the database.

3. PREPROCESSING AND FEATURE EXTRACTION

Fig. 1(a) shows an example of a distance image in false color representation and Fig. 1(b) shows an intensity image of the same scene which was taken with a high resolution 2D camera to provide a reference. Before occupant classification can be performed, some preprocessing is required. In the present paper, only the data from the 3D image sequences are used for the detection and classification task.



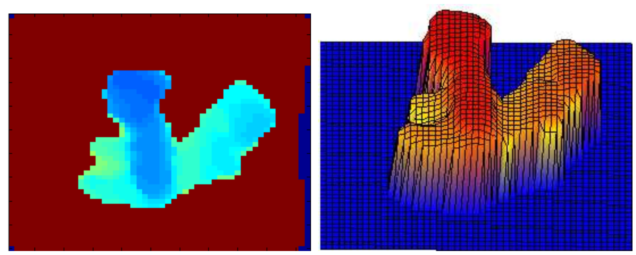
(a) Distance image (b) Intensity image

Fig. 1. Sample distance and intensity image

3.1. Preprocessing

Prior to feature extraction, the image is preprocessed. This involves a distance clipping of the range images; with this operation, range measurements are compared at each pixel location with a reference distance image that corresponds to the empty car interior. This allows to remove any information regarding the background (or objects outside the car), i.e. a binary image can be generated where all background pixels are set to 0 and non-background pixels to 1 (see also [2]).

Fig. 2(a) shows the preprocessed image of the example shown in Fig. 1(a) and Fig. 2(b) shows the topographical view of the same image. Finally, a 2D median filter is applied in order to reduce the range measurement noise present in the image.



(a) Post preprocessing (b) Topographical view

Fig. 2. After preprocessing and its topographical view

3.2. Feature Extraction

Feature extraction involves feature computation and feature subset selection.

3.2.1. Feature Computation

Feature computation aims at obtaining a compact representation of significant information required to describe the relevant parts of the original image. The goal is to preserve as much classification information as possible contained in the original images. We use descriptors that are either derived from the range frame itself or from the representation of the data in the cartesian vehicle coordinate system. Shape features can be calculated directly from a binary 2D range image. By keeping only pixels in the vicinity of a discontinuity in range, an edge image can be calculated, for which contour descriptors can be derived, e.g. area, height and orientation of ellipsoidal contours. Additional features can be gained from the distribution of scatter points in the 3D vehicle coordinate system. Therefore, the coordinates are projected on certain planes and then fitted to different shapes like ellipses or planes. From the fitted shapes we can gain the information about the object for example size, height, volume etc.. In total around 37 features are extracted, which form the input of the classification algorithm.

3.2.2. *Feature Subset Selection*

The computed features may contain redundant information. It is desirable to reduce the size of the feature set to gain robustness in classification performance. Feature subset selection aims at evaluating the effectiveness of individual features or their combination for classification, and selects only the effective ones. This requires an evaluation criterion and a search algorithm. The evaluation criterion evaluates the capacity of the feature subsets to distinguish one class from another or the classification accuracy, while the search algorithm explores the potential solution space. Herein *Sequential Forward Selection* (SFS) and *Sequential Backward Elimination* (SBE) search methods are evaluated as search algorithms to select the feature subset. The mean residual error and Mahalanobis distance are used as a selection criteria[5].

4. CLASSIFICATION TASK

4.1. Bayes Quadratic Classifier

The structure of Bayes quadratic classifier is determined by the conditional densities $p(v|\omega_i)$ as well as by the prior probabilities $P(\omega_i)$. In pattern recognition applications, we rarely, if ever, have this kind of complete knowledge about the apriori probabilities $P(\omega_i)$ and class conditional densities $p(v|\omega_i)$. In the present study, all classes are assumed to be are equiprobable.

The conditional probabilities can be estimated by assuming a multi-variate normal distribution. This is given by,

$$\begin{aligned} p(v|\omega) &= N(v,\mu,\Sigma) \\ &= \frac{1}{\sqrt{(2\pi)^N|\Sigma|}} \exp\left[-\frac{1}{2}(v-\mu)^T\Sigma^{-1}(v-\mu)\right] \end{aligned} \quad (1)$$

where μ is the mean and Σ is the covariance matrix of the input data v.

Now the probability density function of finding a pattern is given by (Bayes Formula),

$$P(\omega_j|v) = \frac{p(v|\omega_j)P(\omega_j)}{p(v)} \quad (2)$$

where,

$$p(v) = \sum_{j=1}^{c} p(v|\omega_j)P(\omega_j)\ . \quad (3)$$

4.2. Polynomial Classifier

Classifiers based on polynomial regression are well-known techniques [4, 6]. The advantage with this approach is that it makes no assumptions about the underlying statistical distributions and leads, at least when using the least mean-square error optimization criterion, to a closed solution of the optimization problem without iterations.

Here, the discriminant function is given by

$$d(v) = A^T x(v) \quad (4)$$

where $x(v)$ is a vector of polynomial terms of an input feature vector v. A is a coefficient matrix which is to be optimized using training samples and is given by

$$A = E\{xx^T\}^{-1}E\{xy^T\} \quad (5)$$

where $E\{\}$ denotes the expecation value taken over all training data. A detailed description of the polynomial classifier can be found in [4].

The discriminant function has as many components as there are classes to be discriminated. Finally the decision is based on the nearest neighbor principle,

$$Best\ match = \arg\ \max_i(d_i(v)) \quad (6)$$

5. RESULTS

In this section the two classification approaches discussed in Section 4 are compared. To investigate the significance of the feature subset selection for the classification task, two experiments were conducted. First, the classification results based on the entire feature set are presented. Thereafter the classification results with the smaller number of features, selected by the feature subset algorithms discussed in section 3.2.2, are presented.

Table 1 shows the classification matrix based on the Bayes quadratic classifier for the test data. We have chosen equal a priori probabilities for all classes. The overall performance is calculated by the number of frames correctly classified divided by the total number of frames. The over all performance is 90.66%. Table 2 shows the classification matrix based on the polynomial classifier. The overall performance in this case is 93.62%. Note that the overall error rate decreases by approximately 15% from 9.44% to 6.38% which is indeed a significant improvement. With the basic configuration, the polynomial classifier shows a clear improvement in the performance when compared to the Bayes quadratic classifier.

In the second experiment, feature subset selection is performed and the classification methods are applied for the same data with the small feature set. The results of the feature subset selection algorithms are not presented here.

Estimated Class

True Class	Empty	RF	FF	P
Empty	93.9	0	0	6.1
RF	0	82.2	0	17.8
FF	0	0.4	83.46	16.5
P	0	3.3	0	96.7

Table 1. Bayes classifier results

Estimated Class

True Class	Empty	RF	FF	P
Empty	100	0	0	0
RF	0.4	97.3	0.3	2.0
FF	0	0.4	98.8	1.2
P	0.6	10	1.2	91.2

Table 2. Polynomial classifier results

However the feature subset selection with the *Sequential Forward Selection* (SFS) search method with mean residual error as a criterion displays better results than the other methods which were discussed in Section 3.2.2. We observed that the feature set can be reduced from 37 to 11 features with addition improvement in form of robustness in the overall performance. Table 3 and Table 4 display the classification matrices for the same testing data as above, with the selected feature subset using Bayes quadratic classifier and the polynomial classifier respectively. The overall performance for the Bayes quadratic classifier increased from 90.7% to 92.2% and for the polynomial classifier case it increased from 93.6% to 95.48%. With optimum feature subset, the size of the training set is reduced which increases robustness of the system.

Estimated Class

True Class	Empty	RF	FF	P
Empty	94.2	0	0	5.8
RF	0	92.3	0	7.3
FF	0	0.4	98.3	1.3
P	0	8.3	0.4	91.3

Table 3. Bayes classifier results with 11 features

Estimated Class

True Class	Empty	RF	FF	P
Empty	97.64	0	0	2.36
RF	0	97.90	0	2.10
FF	0	0.13	99.87	0
P	0	7.14	0.39	92.47

Table 4. Polynomial classifier results with 11 features

6. CONCLUSIONS AND FUTURE WORK

This paper discussed and evaluated two types of classification approaches for the occupant position detection in vehicles based on 3D image sequences. Bayes quadratic classifier uses specific statistical model assumptions whereas the polynomial classifier avoids making explicit assumptions on the underlying distribution. We used the same training set and test/validation set for all the experiments in the present study, thus, the results are directly comparable. Our experiments show that with appropriate feature subset selection, the polynomial classifier achieves the best overall performance (95.48%). Furthermore, it has the lowest computational cost for training and testing classifiers.

Although the polynomial classifier performance is promising there is still room for improvement. For certain classes, test data deviates too much from the training data. This discrepancy can be reduced by increasing the amount of training data. However, in real-world environment, all possible situations that may occur in a car can not be considered in the training database. Therefore we need to find a confidence or rejection function that allows suppressing wrong classifications that might occur in such cases. These techniques require further study.

7. REFERENCES

[1] S. Lacroix S. Gautama and M. Devy, "Performance of stereo for occupant detection," *Proceedings of the international workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, RATFG-RTS'99,*, pp. 177–184, 1999.

[2] Martin Fritzsche, Christian Prestele, and Bruno Mirbach, "Vehicle occupancy monitoring with optical range-sensors," *IEEE Intelligent Vehicles Symposium,*, pp. 90–97, 2004.

[3] Duda R. O. and Hart P. E, *Pattern Classification and Scene Analysis*, Wiley, 1991.

[4] Jürgen Schürmann, *Pattern Classification: Statistical and Neural Network based Approach*, John Wiley and Sons, Inc., New York, 1990.

[5] K. Z Mao, "Orhogonal forward selection and backward elimination algorithms for feature subset selection," *IEEE Trans. On Systems, Man, And Cybernetics-Part B: Cybernetics.,*, vol. 34, no. 1, 2004.

[6] W. M. Campbell and K. T. Assaleh, "Polynomial classifier techniques for speaker verification," *IEEE International Conference on Acoustics, Speech and Signal Processing.*, vol. 1, pp. 15–19, 1999.