# Occupant Classification System for Automotive Airbag Suppression

Michael E. Farmer[§] and Anil K. Jain*

[§] Eaton Corporation

* Michigan State University

*Email: farmerm3@msu.edu, jain@cse.msu.edu*

## Abstract

*The introduction of airbags into automobiles has significantly improved the safety of the occupants. Unfortunately, airbags can also cause fatal injuries if the occupant is a child smaller (in weight) than a typical 6 year old. In response to this, The National Highway Transportation and Safety Administration (NHTSA) has mandated that starting in the 2006 model year all automobiles be equipped with an automatic suppression system to detect the presence of a child or infant and suppress the airbag. The classification problem we address is a four-class problem with the classes being rear-facing infant seat, child, adult, and empty seat. We describe a machine vision-based occupant classification system using a single greyscale camera and a digital signal processor that can perform this function in "real time" (< 5 seconds). The system has been extensively tested on a database of over 21,000 real-world images collected over a period of 4 months in moderate lighting conditions with a wide variety of passengers in eight different vehicles. We have achieved a classification accuracy of ~95%. . We believe this system serves the need for a low-cost, high reliability embedded real-time airbag suppression system. Additional testing and improvements of the classification system are currently underway.*

## 1. Introduction

Recently, there has been considerable attention paid to developing 'smart' airbags that can determine not only if they should be deployed in a crash event but also with what force they should be deployed [1, 2, 3, 4]. The size and type of occupant needs to be used to determine the safe level of force used to deploy the airbag.

In May 2001 the U.S National Highway Transportation and Safety Administration (NHTSA) defined the Federal Motor Vehicle Safety Standard (FMVSS) 208 that mandates automatic airbag suppression when an occupant smaller than a 6 year old child is in the passenger seat [8].

The standard also calls for suppression of the airbag for Rear Facing Infant Seat (RFIS) and Forward Facing child seats and booster seats and enabling the airbag when the occupant is a 5th percentile (by weight) adult female or larger. This standard was proposed because a number of small children and infants have been killed by airbags. Between 1986 and 2001, 19 infants in RFISes and 85 children were killed by airbags [9]. Figure 1 shows the effects of an airbag on a RFIS during deployment.

The recognition of the occupant type (i.e., RFIS, child, adult, empty seat) is referred in the NHTSA standard as *static suppression*. The NHTSA specification is written around the use of a seat weight sensor, which measures the weight of the occupant and infers whether it is of sufficient weight to be an adult. A wide variety of systems have been proposed for solving this problem [1, 2, 3, 4]. The aim of this paper is to show the applicability of machine vision (using a single B/W camera) to solve the airbag static suppression problem. We highlight some of the difficulties of the airbag suppression problem, propose a novel processing framework, and show the resulting performance.

## 2. Airbag Suppression Problem

The airbag suppression decision problem can be defined as a 4-class problem: (i) RFIS, (ii) Child (< 25.6 kg in weight & < 124.5 cm tall), (iii) Adult (> 46.7 kg in weight & > 139.7 cm tall), and (iv) Empty. The airbag is only deployed when the passenger is assigned to the adult class. Figure 2 shows examples of each of the classes. According to the NHTSA specification, the system must determine the classification of the occupant within 10 seconds of a change of occupant state (i.e., from empty to adult) [8]. In addition to this time constraint, the NHTSA specification also requires 100 percent correct classification for a subset of possible seating positions for all of the occupant types listed above.

The use of computer vision in the automobile environment is challenging due to the extreme variations in lighting from bright daylight to dark night as seen in Figure 3. Additionally, in very bright sunlight the image

may have considerable dynamic range due to the simultaneous existence of shadows near the occupant's legs and bright patches due to direct sunlight on the head and torso. Since the vehicle is moving, there are both moving and stationary shadows caused by sunlight that further complicate the problem.



**(a)**          **(b)**

**Figure 1. RFIS during airbag deployment; (a) RFIS before deployment, (b) RFIS after deployment.**



(a)         (b)

(c)         (d)
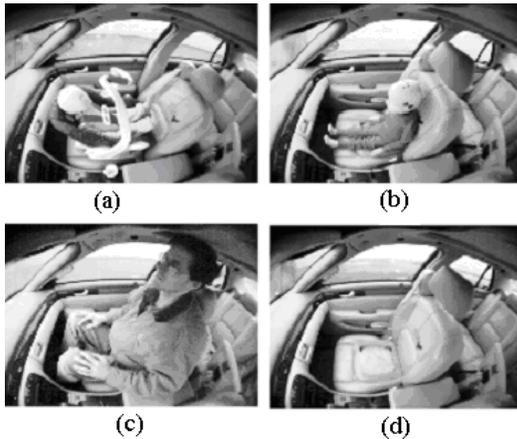
**Figure 2. Examples of the four classes; (a) RFIS, (b) child, (c) adult, (d) empty.**



**Figure 3. Demonstration of lighting variability.**

Other complications include the large intra-class variability for three of the classes mentioned above (the empty seat class has very little intra-class variability aside from lighting changes). For the child and RFIS classes,

there are a number of seat types as well as seating positions that must be recognized, and the similarity between them is often not very high as shown in Figure 4. One further complication is that RFIS and booster seats may be covered with blankets to protect the children from sunlight and cold (see Figure 5) [8]. The adult class also has a large amount of intra-class variability, as shown in Figure 6, due to the following three factors:

1) Variability from the $5^{th}$ percentile female to the $95^{th}$ percentile male is 10 inches and 75 pounds.
2) Variability in adult appearance due to hair and clothing variations.
3) Seasonal variability as clothing changes from only shirts to down parkas and hats. This variation is present not only from person-to-person but for the same person from season-to-season.

To summarize, a vision-based system for airbag suppression must be robust enough to handle the following conditions:

1) Large intra-class variability of the four classes
2) Camouflaged classes (e.g., blanketed RFIS)
3) Large variation in light levels (day to night)
4) Large lighting variations within an image (shadows to bright direct sunlight)
5) Severe automotive environmental conditions
6) Low cost
7) Extremely high reliability and performance

## 3. System Description

Figure 7 shows the block diagram of the proposed airbag suppression system. The classification processing is executed every 3-5 seconds to meet the NHTSA specification [8]. The system is physically composed of a single monochrome digital CMOS camera with a wide field-of-view lens, a bank of LED illuminators, a Digital Signal Processor, and a control micro-processor. The camera is a commercial-off-the-shelf 400x320 pixels camera without the standard IR filter to allow the camera to use the supplemental IR illumination in dark conditions. The illuminators consist of a bank of infra-red LEDs geometrically configured to provide roughly uniform illumination over the passenger area of the vehicle. The illuminators also contain a diffuser to ensure eye safety and uniform light levels over the field-of-view.

The Digital Signal Processor performs the image processing functions in real-time. The micro-processor is responsible for system diagnostics and for maintaining communications with the other subsystems in the vehicle via the vehicle bus and for providing the suppression signal to the vehicle airbag control module.

The system is located in the roof liner of the vehicle along the vehicle center-line and near the edge of the windshield. This location provides a near profile view of

the occupant in the passenger seat, which aids in the classification of the occupant. This location also reduces the likelihood of the occupant blocking the sensor and makes styling the sensor into the vehicle easier. The typical field of view required for most passenger vehicles is roughly 100 degrees vertical FOV and 120-130 degrees horizontal FOV. This FOV ensures coverage of the occupant from the Instrument Panel to the rear-most seating position when the seat is fully reclined.



**Figure 4. Intra-class variability for RFIS and child classes.**



| (a) | (b) |
|---|---|

**Figure 5. RFIS (a) and RFIS under blanket (b).**

The first stage in the Static Suppression (Classification) Processing is the Segmentation (see Figure 7). Due to space constraints, we are unable to provide details of the segmentation algorithm. Upon completion of the segmentation, shape features are extracted from the resultant region of interest. There are numerous methods for computing features of objects where shape characteristics are considered important [11].

For images with reasonable lighting, we use edge-based features while for night-time imagery where the contrast is considerably lower, we rely on silhouette features. The edge features consist of the complete internal edges of the occupant and the seat simultaneously as shown in Figure 8. Notice that the child is nearly completely engulfed inside the seat boundary. Therefore,

a traditional contour/boundary edge following algorithm such as Fourier Descriptors would be incapable of recognizing the difference between a child class and the empty class [11].



**Figure 6. Intra-class variability for the adult class.**
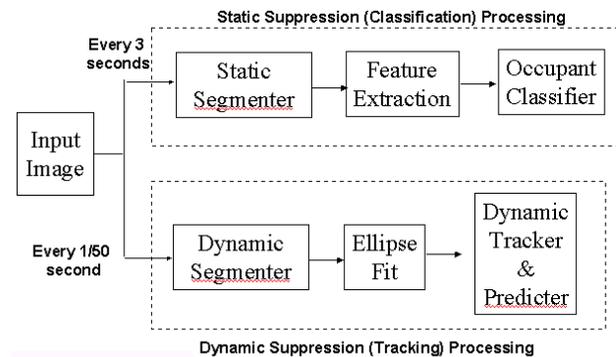


**Figure 7. Diagram for the Airbag Suppression System.**

Due to the severe lighting conditions where one part of the image is very dark and another is very bright, a simple adaptive threshold detector would often miss edges in an entire region of the image where the contrast is too

low due to either saturation or low light levels. Also the edges tend to be heavily clustered into regions of high contrast when a simple adaptive threshold technique is used. We have found however, that the adaptive threshold edges tend to be more sensitive at distinguishing RFIS class so it is used for this purpose.
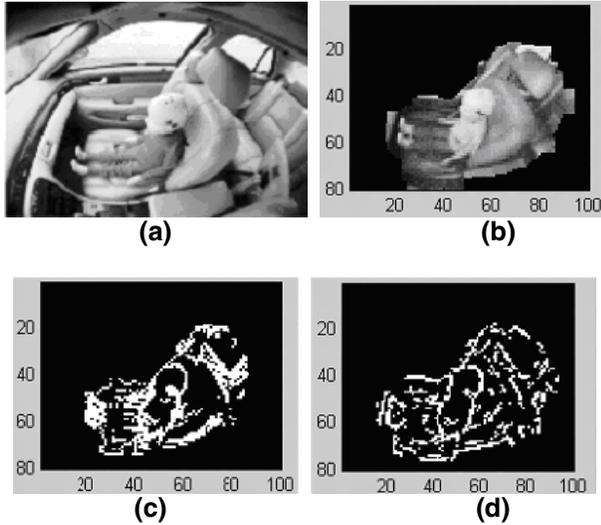


**Figure 8. Edge detection (a) input image, (b) segmented image, (c) adaptive threshold edge detection results, (d) CFAR detection.**

To maintain sensitivity for the other classes, we compute edges based on another edge detection process. The first stage processes the image with a simple gradient calculator, generating the x and y directional gradient magnitudes at each pixel. The edge magnitude is then computed and processed with a Constant False Alarm Rate (CFAR) based detector. Figure 8 shows the resultant edge map for the traditional edge detection and for the CFAR detection [10].

We used the Cell-Averaging CFAR where the average edge amplitude in the background window is computed and compared to the current edge image as shown in Figure 9 [10]. Only the pixels that are non-zero are used in the background window average. The guard region separates the pixel of interest from the background window. For the results in this paper, a 5x5 CFAR kernel (see Figure 9) is used. A pixel is marked as an edge pixel if the ratio of the test sample amplitude to the background region statistic exceeds a threshold that is set by the desired false alarm rate of the detector).

The CFAR Processing generates a relatively uniform distribution of edges across the image. In order for the system to properly classify the occupant, the small, disconnected edge pixels are removed. This leaves the major edge elements such as occupant boundary along the seat which are effective for discriminating occupant types.

The mechanism for removing the smaller edges is to discard edges below a certain spatial extent.
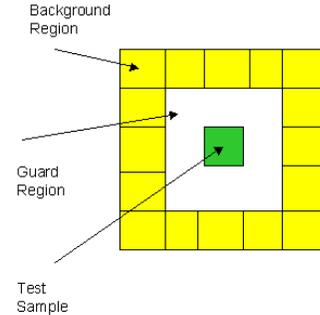


**Figure 9. CFAR edge detector.**

Once the edge image has been obtained, the feature extraction processing is performed. There are numerous candidate features that can be used [11]. The feature set we chose is the Legendre Moments representation of the edge image. Both Chebyshev and Legendre moments were considered due to their efficient representation of an image based on a finite set of moment values [7]. Legendre moments were chosen over Chebyshev moments due to the slightly simpler generating function and only a slight advantage in reconstruction accuracy [7]. The recurrence relation for the one-dimensional Legendre Polynomials is [6]:

$$p(0) = 1; \quad p(1) = x;$$
$$p(n) = 1/n * ((2*n-1)*x*p(n-1) - (n-1)*p(n-2)).$$

. While the above relation is for one dimension, the two-dimensional representation is simply a product of the two dimensions. The generating function for the Legendre moments from geometric moments is [6]:

$$L_{mn} = \frac{(2m+1)(2n+1)}{4} \sum_{j=0}^{m} \sum_{k=0}^{n} C_{mj} C_{nk} M_{jk},$$

where $C_{ij}$ are the coefficients defined by the above polynomial generating function and the $M_{ij}$ are the traditional geometric moments.

We have experimented with the effects of moment order on classification accuracy, and these results are provided in Section 4. The qualitative effects of moment order can be seen in Figures 10 and 11. Figure 10 shows the segmentation and the edge images for a RFIS and an adult. Figure 11 shows the resultant reconstructions using moments up to the $26^{th}$, $36^{th}$, and $45^{th}$ order. Clearly, there is a significant improvement in representation as the moment order is increased.

Note that the complete set of moments up to the $45^{th}$ order generate 1,081 features. To avoid the "curse of dimensionality", we apply a feature selection algorithm to reduce the feature set to 50 features for each class. With 4 subclasses (RFIS, child, adult, empty), this generates a

maximum of 200 features. Since there are often some common features between the classes, the final number of features is roughly 150. The feature selection algorithm is a filter method where we rank the relative discriminative ability of each feature and select the best features. The measure of discrimination is based on the Mann-Whitney test statistic. In addition to feature selection, the features are also scaled to be between [0, 1] to reduce the effects of large amplitude features skewing the distance metrics of the classifier.
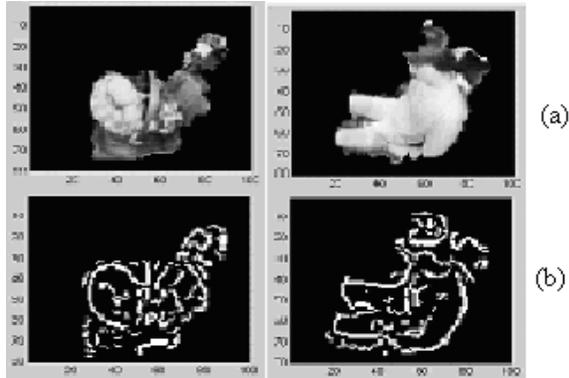


**Figure 10. Segmented and edge image for RFIS (a) and Adult (b).**

The other two paths are for the two types of k-NN classifiers: the traditional k-NN classifier and a distance-based k-NN classifier. For both of these classifiers, k=9 provided the best results. The distance-based k-NN classifier computes the average distance of the test sample to the k-nearest training samples in each class. For example, it computes the mean for the top k RFIS training samples, the top k adult samples, etc. The final decision is to choose the class with the lowest average distance to its k-nearest neighbors. The distance metric we use is the Manhattan distance for each of these classifiers. The classifier combiner shown in Figure 12 simply takes an average of the probabilities from each of the four classifiers.

The classifier uses four different nearest neighbor-based classifiers as shown in Figure 12. The two paths shown correspond to the types of edges used to generate the moment features (traditional adaptive threshold edges and CFAR edges).

## 4. Experimental Results

The airbag suppression system has been integrated into an automobile for both training and test data collection. Performance testing of the classification system has been done using the Leave-One-Out method. The test database consists of 21,000 images collected over a 4 month period. The data was collected in 5 different cars over a period of 4 months. The data consisted of 20 different

types of child seats. It also consisted of both ATDs (crash test dummies) and real adults. The breakdown of the imagery is as follows:

| RFIS | 7753 |
|---|---|
| child | 5994 |
| adult | 3243 |
| - ATD | 619 |
| - human | 2624 |
| Empty | 4868 |
| | |
| Total | 21858 |

The RFIS and forward facing child seat images were collected with realistic looking dolls used as subjects. The three-year old and six-year old subjects were (ATDs). Table 1 provides the confusion matrix for testing against the entire database using maximum moments of orders 26, 36, and 45. An additional larger database will be used for future testing that contains data collected at local schools to enable us to collect human child images. The current images were collected in indoor and outdoor conditions, however the vehicle was not moving due to collection difficulties with moving vehicles.

The classification results are encouraging considering the large intraclass variation with an overall classification accuracy of better than 95 %. The expected performance for this system is very aggressive. The NHTSA specification mandates that the system perform with 100% accuracy on a set of well-defined test conditions [8]. Unfortunately, these test conditions mandate the covering of all of the infant carseats with large blankets. This causes overlap into the adult class that may not be repairable. The test database did not contain blankets.

These current results are the output of an ongoing research and development project, and we will continue to release additional results as they become available. An example of an image that was mis-classified can be fond in Figure 13. This image was a child but was mis-classified as an empty seat due to the poor segmentation.

## 5. Summary and Conclusions

We have shown the applicability of machine vision to the airbag static suppression problem. We have highlighted some of the challenges in designing a machine vision system for this application domain. A novel processing framework using two parallel edge detectors for feature extraction is proposed. A classifier combination strategy based on multiple k-nearest neighbor classifiers is defined. Overall, the system provided a correct classification rate of 95 % on a database of over 21,000 images collected in an automobile in a variety of occupant seating scenarios.

Our future work will be directed at testing in actual drive conditions where the history processing can be fully tested. More robust combination strategies will also be investigated for use in these drive conditions. We will also investigate using even higher precision math to allow the use of higher order moments. We will also continue to improve the edge detection processing to ensure the system operates in even the most severe lighting conditions. The moments of these edge images are also only one possible feature set. We will investigate using a combination of these features with additional shape features to seek further improvements.
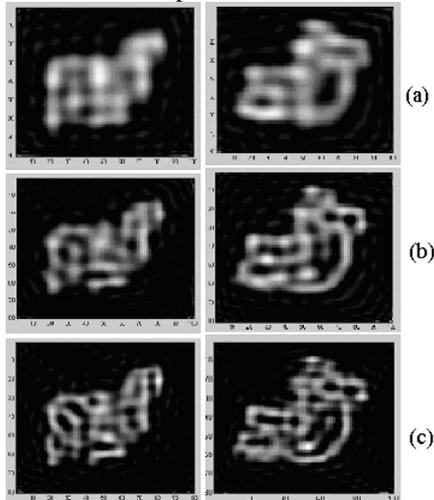


**Figure 11. RFIS and Adult reconstructions for moments of up to 25th (a), 35th (b), and 45th (c) order.**
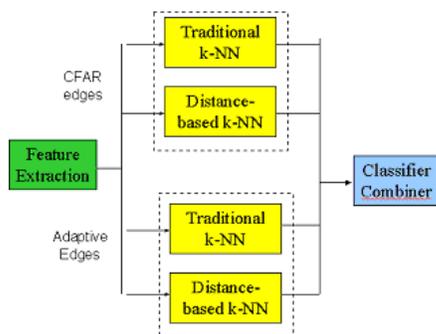


**Figure 12. Classifier architecture.**

(a)    45th order moments

|       | RFIS | child | adult | empty |
|-------|------|-------|-------|-------|
| RFIS  | 7053 | 590   | 5     | 1     |
| child | 13   | 5819  | 2     | 5     |
| adult | 19   | 490   | 2615  | 16    |
| empty | 0    | 12    | 0     | 4830  |

(b)    36th order moments

|       | RFIS | child | adult | empty |
|-------|------|-------|-------|-------|

| RFIS  | 6986 | 657  | 5    | 1    |
|-------|------|------|------|------|
| Child | 15   | 5813 | 6    | 5    |
| Adult | 18   | 488  | 2619 | 15   |
| Empty | 0    | 14   | 0    | 4828 |

(c)    26th order moments

|       | RFIS | child | adult | empty |
|-------|------|-------|-------|-------|
| RFIS  | 6942 | 698   | 7     | 2     |
| child | 7    | 5820  | 3     | 9     |
| adult | 21   | 530   | 2570  | 19    |
| empty | 0    | 17    | 0     | 4825  |

**Table 1. Leave-one-out classification results.**
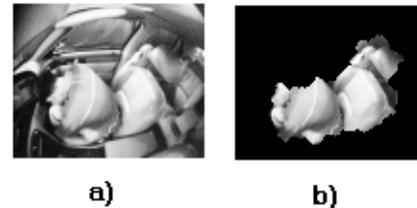


**Figure 13. Classification error example (a) original image of adult, (b) segmented image thought to be a RFIS.**

## References

[1]   M. Farmer, R. Hsu, and A.K. Jain, "Interacting Multiple Model (IMM) Kalman Filters for Robust High Speed Human Motion Tracking", 16th Intl. Conf on Pattern Recog., Vol 2, pp 20-23, 2002.

[2]   P. Mengel, G. Doemens, and L. Listl, "Fast range imaging by CMOS sensor array through multiple double short time integration (MDSI)", *Proc. ICIP*, pp. 169-172, 2001.

[3]   A.P. Corrado, S. Decker, and P. Benbow, "Automotive occupant sensor system and method of operation by sensor fusion", *US Patent 5482314*.

[4]   J.H. Semchena, E. Faigle, R. Thompson, J. Mazur, and C. Steffens Jr., "Apparatus and method for controlling an occupant restraint system", *US Patent 5531472*.

[5]   J. Krumm and G. Kirk, "Video occupant detection for airbag deployment," *Proc. IEEE Workshop on Applications of Computer Vision*, pp. 30-35, 1998.

[6]   M.R. Teague, "Image analysis via the general theory of moments", J. Opt. Soc. Amer. Vol. 70, No. 8, pp. 920-930, 1980.

[7]   R. Mukundan, S.H. Ong, and P.A. Lee, "Image Analysis by Tchebichef Moments", IEEE Transactions on Image Processing, Vol. 10, No. 9, pp. 1357-1364, 2001.

[8]   National Highway Transportation & Safety Administration, Federal Motor Vehicle Safety Standard # 208, 2001.

[9]   General Accounting Office, Vehicle Safety - Technologies, Challenges, and Research and Development Expenditures for Advanced Air Bags, June 2001.

[10] H. Vincent Poor, An Introduction to Signal Detection and Estimation, Springer, 1994.

[11] Oivind Due Trier, Anil K. Jain and Torfinn Taxt, Feature Extraction Methods for Character Recognition – A Survey, Pattern Recognition, Vol. 29, No. 4, pp. 641-662, 1996.