

## EMERGING CCD-CMOS TECHNOLOGY: AN OPPORTUNITY FOR ADVANCED IN VEHICLE SAFETY SENSORS

**Laurent Federspiel**

IEE S.A.

Luxembourg

Paper Number 05-0116

### ABSTRACT

Passenger airbags are currently designed for the optimal support of a 50-percentile adult in a crash, reducing the risk of severe injury for a maximum range of occupants. However, such a fixed-level, high-energy airbag deployment can be extremely dangerous for very small occupants, for example the 5-percentile woman or children in infant seats. For this very reason, new standards such as FMVSS 208 (Federal Motor Vehicle Safety Standard No. 208) include differentiated airbag deployment strategies according to occupant classification.

IEE, Luxemburg, develops and manufactures such occupant classification systems. An example of which are the sensor mats made by IEE, which tier one automotive suppliers use globally for their seating systems. These mats measure the two-dimensional pressure profile in the seat area, and deliver these values for a pattern recognition algorithm as basis for occupant classification. An innovative development project, currently being conducted by the company, is an optical system which can provide three-dimensional information on the occupant, enabling highly differentiated classification. This system is projected to become commercial by 2007.

### LEGAL AND SENSOR REQUIREMENTS FMVSS 208

According to FMVSS 208, restraint systems have to be designed in such a way that, in the event of an impact, they create less risk of airbag induced injuries, particularly for small women and young children.

To achieve these goals FMVSS 208 proposes three airbag deployment strategies in the event of a crash:

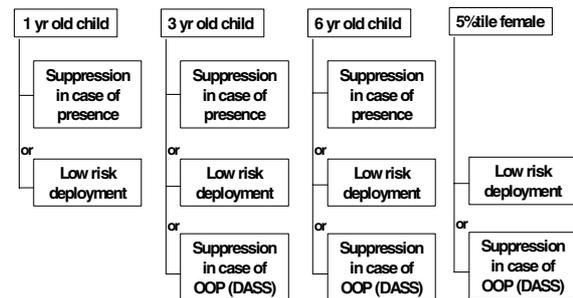
'Suppression In Case Of Presence', if sensors detect an infant seat, occupied by a child up to six years old, and deployment of the airbag, if a person in the range of a 5-percentile woman or taller is detected,

'Low Risk Deployment' (LRD) means that the airbag deployment does not harm an occupant at close range from the airbag module. For verification a dummy is positioned close to the dashboard while the airbag is deployed and the

corresponding dummy injury criteria must not bypass certain values to be in line with the low risk deployment strategy. Sensing technology can be used to switch the airbag to a low output mode.

'Dynamic Automatic Suppression Strategy' (DASS), meaning that in addition to a qualitative occupant classification (as with LRD), the occupant's current position in relation to the airbag deployment door (in-position, out-of-position) has to be traced and the airbag is suppressed if the occupant is at close range to the airbag deployment door.

'FULL' LRD, as well as the sophisticated DASS strategy, require highly-sensitive, advanced occupant classification systems, which can deliver the complex data set necessary to take the best possible decision.



**Figure 1: Different certification strategies proposed by the NHTSA FMVSS208 final rulemaking**

For the standards '3 year-old child', '6 year-old child' and '5-percentile woman', the LRD strategy is already widely used. For the 1 year-old child in a rear-facing infant seat (RFIS) and placed on the passenger seat, both 'Suppression' and LRD are also included in the standards. However, airbag technology does not yet permit energy limitation, as required by LRD. In such cases, today's systems are changed to a controlled switching off of the airbag. This ensures at least a certain minimal security for all accident scenarios.

Dynamic Automatic Suppression Strategy (DASS) provides considerably more opportunity. The newly-developed IEE 3D-System provides the necessary information for differentiated recognition. In all cases of occupation, including the RFIS and out-of-position occupation, this

system provides the essentials to adapt correspondingly-modified airbag modules to the 'Low Output Mode'.

Although a DASS strategy for the 1 year old child is not yet approved and optimized, an airbag strategy for real life child seat scenarios could be as follows:

- an RFIS is always considered as 'out-of-position',
- for a FFCS a specific airbag suppression zone (ASZ) could be defined. Only if the child is out of this area, the airbag will be deployed (with less energy).

Accordingly, the 3D system allows an airbag strategy matched to the situation (RFIS / FFCS / person OOP / person in position), rather than the presently insufficient differentiated strategy based on age. Suggestions for respective test procedures have been submitted for assessment by the US NHTSA (National Highway Traffic Safety Administration).

Requirements for the specific sensors may distinctly differ, depending greatly on the OEM's own safety strategy and the individual design of the car (small roadster or large truck). On the other hand, the installed sensor families should meet differing safety requirements in the US and other parts of the world.

## RELATED WORK

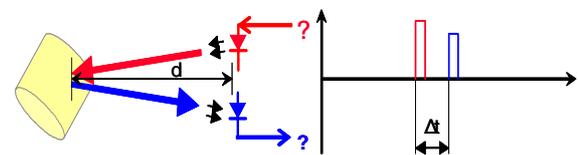
Current technology for occupant classification systems is mainly based on pressure sensors installed inside the vehicle seats, for e.g. , the IEE OC<sup>®</sup> sensor. However, with vision based system, the position of the occupant and orientation of the child seat can be also be determined. Different approaches in the vision systems can be broadly divided into different categories based on the sensing technology. Sensing technology is either based on the video camera (for example see [1]) or on stereo-vision based range images ([2], [1], [3], [4]). In the reference [2], a 3-D vision system using stereo cameras was developed. It was argued that stereo vision offers a potential to produce detailed results within real-time constraints and it suited for irregular environment. In references [3] and [4], stereo-based range data was used to detect whether and where humans are inside a vehicle. In [1], Krumm and Kirk developed a system based on both intensity (2-D) and stereovision-based (2 and half-D) range data and found for each class the principle components, with which nearest neighbor classification was performed. However, these methods are based on stereo vision which are sensitive to varying illumination conditions inside the car. Furthermore, extra equipment and processing is required to capture 3-D information from the stereo images. Another important aspect for a serial production is

the cost of such a system. Hence, above systems are definitely not cost effective as they require two cameras for capturing the scene, and the need of important processing power and time.

## REAL TIME 3D TIME-OF-FLIGHT IMAGING

Key element of the new optical occupant classification system developed by IEE is a 3D Modulated Light Intensity (MLI) System. The system's ability to deliver three dimensional images is based on the measurement of the phase shifts of the modulated emitted light signal and its reflection by the object. The smaller the difference, the shorter the distance between the object (the occupant or the infant seat) and the sender/recorder-combination. Thus every snap-shot delivers an image with differentiated depth information for the complete detection area.

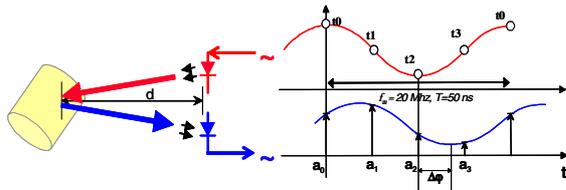
Other time of flight (TOF) technologies apply a different principle emitting a short pulse of high optical intensity (Figure 2). The light velocity turns into a flight time of only 66ps per meter distance (resolution 1cm). These short periods require sensors of extremely high sensitivity. In order to obtain a resolution in the 1 cm range, the frequency bandwidth has to be greater than 10 GHz. This in turn creates high energy consumption, which is difficult to supply in the automotive industry.



**Figure 2: Light pulse based time of flight. The turn around time of an emitted light pulse is measured and put into relation of the distance  $d=c \cdot \Delta t / 2$**

The IEE MLI System uses a different approach. By emitting a continuous wave-modulated cone of light, with a defined wave length, the phase difference between sent and detected signal can be measured and to generate a topographic image provided afterwards to the classification algorithm (Figure 3). This principle, which consumes much less energy, is the basis on which the IEE system works.

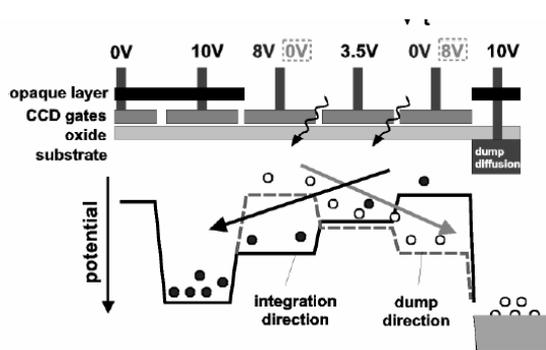
A key feature is an active, non-scanning light source, which emits amplitude modulated near infrared light (NIR) and thus delivers a homogeneous illumination for the camera field of vision (FOV).



**Figure 3: Detected light intensity as a function of time. The sinusoidal modulation (top curve) of the illumination causes a periodically modulated signal in the receiver (lower area). The phase offset can be computed by evaluating the signal amplitudes  $a_0, \dots, a_3$  at 4 different temporal positions  $t_0, \dots, t_3$ .**

Due to the travel time of the light to and from the target, the phase of the detected beam is retarded compared to the phase of the modulation signal in the transmitter (see Fig. 3). This phase delay can be measured and directly converted into the distance between the target and the camera. The amplitude and phase of the received signal can be retrieved by synchronously demodulating the incoming modulated light within the detector. Demodulation of a received modulated signal can be performed by correlation with the original modulation signal (cross-correlation). The measurement of the cross-correlation function at selectively chosen temporal positions (phases) allows the phase of the investigated periodical signal to be determined [5]. With the selected temporal positions  $t_0 = 0^\circ, t_1 = 90^\circ, t_2 = 180^\circ, t_3 = 270^\circ$ , one can calculate the phase offset via the formula

$$A = \arctan \left[ \frac{a_3 - a_1}{a_2 - a_0} \right] \quad (1)$$



**Figure 4: Cross sectional view of the CCD pixel layout**

Figure 4 shows the layout and a cross-section view of the pixel. By applying proper gate voltages to the photo gates, the potential gradient in the semiconductor is influenced. If the voltages of the photo gates are changed synchronously with the modulated light, optically generated charge carriers

move either to the integration gate (IG) or are dumped to the dump diffusion. This process is repeated until the integration gate has accumulated a sufficiently large signal. The four amplitudes  $a_0, \dots, a_3$  are obtained by subsequently repeating this process at 4 different phase offsets [5]. The IEE sensor is based on a 4 tap-pixel sensor, a design which acquires the 4 amplitudes simultaneously.

With regards to system accuracy, the assumption is made that depth is not limited by electronics/noise of the detection system but only by the photon shot noise (a physical limit). Achieved accuracy can therefore be calculated, and depends on

- background illumination and other noise sources, and
- on the object reflectance and its distance to the sensor.

The dependence of reflectance and background noise is calculated and read out as relative fault of the amplitude value. This ensures adequate action can be taken should measurement error become too great.

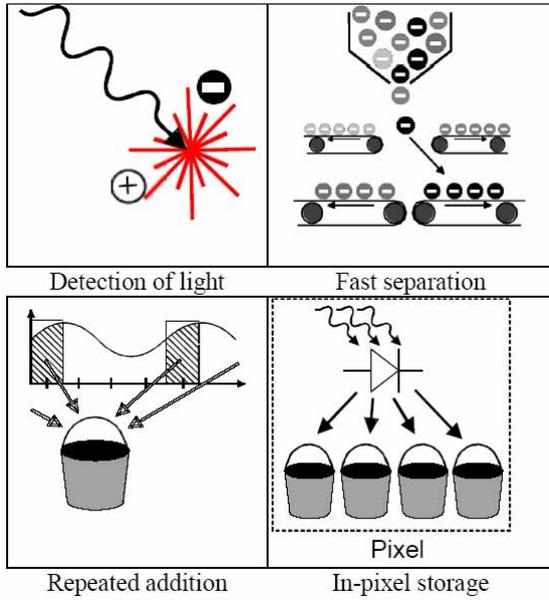
Moreover, the mean amplitude value per pixel (corresponding to the intensity of the reflected light) allows the generation of a grey scale image of the complete detection area.

In summary, the main advantages of the IEE 3D-Camera-Solution are the simultaneous provision of distance information and accuracy, combined with a real life b/w image.

## CAMERA HARDWARE

A monocular camera is integrated in the vehicle's center overhead module, enabling a field of view of  $120^\circ \times 90^\circ$  ( $136^\circ$  in the diagonal) with a resolution of 50 by 52 pixels. Using a near infrared light, unperilous to the human eye, at a wavelength of around 890nm, sensing range is up to 750cm (limited by the modulation frequency of the light). At a distance of 150cm, depth accuracy is at 2.2cm.

The sensor as key component of the whole system is realized in a 4-tap pixel architecture. This 4-tap pixel is built in form of two single 2-tap structures. These two structures are controlled in a way that the phases  $0^\circ$  and  $180^\circ$  as well as the phases  $90^\circ$  and  $270^\circ$  can be captured in parallel.



**Figure 5: Key function of the imager: Detection of light, fast separation of the generated electrons into the 4 different taps; repetition of the measurement until reliable signal generated and storage in of signal in pixel before reading out complete imager**

The sensor transforms the incoming optical signal into electron-hole pairs. The efficiency of this process is basically limited by the inherent quantum efficiency of the chosen semiconductor material and the fill-factor of the optical sensor. In order to demodulate the incoming 20MHz signal, a fast charge separation and transport has to take place within each pixel. The sensor's ability to separate and transfer the charges to the corresponding output node represents the demodulation contrast (2), which is defined as the ratio of the demodulated amplitude A (1) and the acquired offset signal B,

$$C_{\text{demodulation}} = \frac{A}{B} \quad (2)$$

Within one single modulation period of 50 ns (corresponds to the modulation frequency of 20 MHz) typically only a few photons impinge on each individual pixel and hence only a few photoelectrons are generated in the pixel. For a broad range of operating conditions – statistically spoken – even less than one electron is generated per modulation period. The repeated addition of the electrons generated over numerous modulation periods is thus necessary and represents a very important feature of the current embodiment. The approach of adding charges almost noise-free at the pixel level is tightly linked to the CCD pixel realized in a CMOS technology. This CCD pixel represents a key element to the success of the

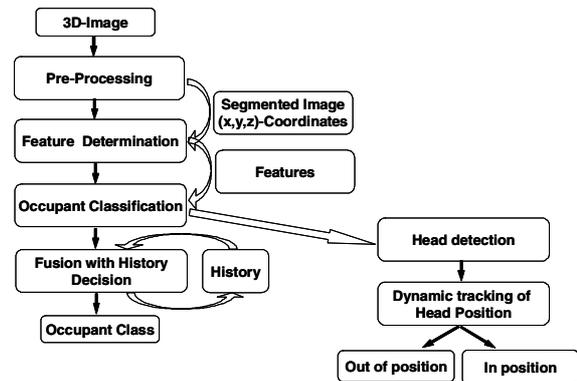
present technique. Moreover, the in-pixel storage and the processing of the different signal samples allow a high degree of flexibility in the readout process.

An automotive occupant monitoring system requires the development of a specific lens for the imager. The optical field of view for an occupant classification system must have an opening of at least 120° in the horizontal x-axis of the vehicle. The point spread function, a low f-number and an application specific anti-reflection coating are only some of the elements which characterizes this lens development.

The active light emitter is realized on a single board. The module is built in a chip-on-board (COB) technology. The illumination unit is covered by a structured lens in order to distribute uniformly and to guide the optical power to the regions of interest defined by the type of the application. The lens provides an additional safety margin to the requirements of the eye-safety norm EN 60851 class 1. The developed system emits a sinusoidal wave illumination front with a total mean power of 600mW.

## STATIC CLASSIFICATION ALGORITHM

The algorithm related to the *static* occupant classification is a three step process (Figure 6): The pre-processing of the data recorded by the camera is followed by a feature determination step and the classification step. In a fourth step, the localization and the *dynamic* tracking of the occupant's head position complete the process

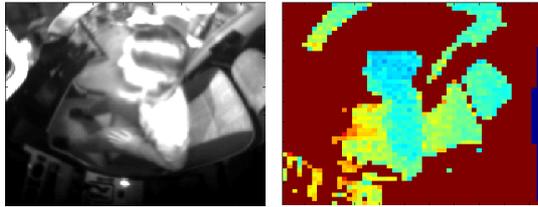


**Figure 6: Algorithm flowchart for static and dynamic classification**

### Step 1: Preprocessing

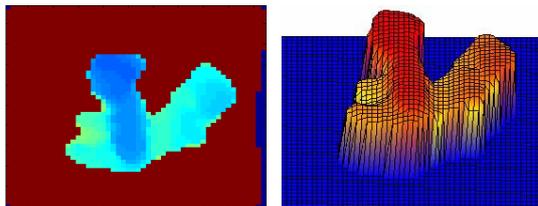
Step one starts with a pre-processing algorithm to reduce the image noise and to eliminate the background. This involves a distance clipping of the range images; with this operation, range measurements are compared at each pixel location with a reference distance image that

corresponds to the empty car interior. This allows removing 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. Once this is done they are then transposed as a three dimensional matrix in a Cartesian coordinate system (Figure 8).



**Figure 7: (left side) intensity image of the scene. (right side) color coded distance raw image before preprocessing**

As the comparison with the inserted b/w image in Figure 7 and 8 proves, multiple information is available about the occupant's head, hands and shoulder position, as well as the occupant's position in relation to the car seat backrest.



**Figure 8: (left) preprocessed distance image; (right) topographical view**

## Step 2: Feature Computation

The second step covers the feature determination. For this purpose the recorded patterns are compared with basic patterns stored in a database, and subsequently characteristic analogies are used to characterize the content of the recorded image. As an example, the comparison of the patterns of an RFIS and a small adult who is sat upright uses indicators like the angles between the typically fixed structures (seat and backrest) and the variable structures, determined by the characteristic seat occupation (slope of the infant seat backrest, position of the person in relation to the car seat backrest). Once the differentiation between infant seat and person has been completed, and a person has been identified, the position of the person's head is detectable.

Feature computation aims in 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 image. This representation in terms of features should be computationally inexpensive so as to fulfill the real-time requirement. Descriptors are used 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 information are gained about the object for example its size, height, volume etc.. In total ten features are extracted, which basically establish the input of the classification algorithm.

**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. Sequential Forward Selection (SFS) search methods are used as search algorithms to select the feature subset. [6].

## Step 3: Classification Method

Step 3 covers the action to be taken in the event of an accident, determined by situation and according to FMVSS 208 LRD. The system has to find out

- if the seat is occupied or not,
- if yes, if the seat is occupied by a FFIS or a RFIS, or
- if the seat is occupied by a small person, the pattern of which corresponds to either a 3 year-old child, a 6 year-old child, or a 5-percentile woman.

A polynomial classifier has been selected for the classification task. Classifiers based on polynomial regression are confirmed techniques [7]. The advantage with this approach is that it makes no assumptions about the underlying statistical distributions and leads, when using the least mean square error optimization criterion, to a closed solution of the optimization problem without iterations.

The discriminate function is given by,

$$d(v) = A^T x(v) \quad (3)$$

where  $A$  is a coefficient matrix which is to be optimized using training samples and is given by,

$$A = E\{xx^T\}^{-1} E\{xy^T\} \quad (4)$$

and  $x(v)$  is the matrix of polynomials of the input feature vectors [7]. The discriminate function has as many components as there are classes defined to be discriminated. Finally the decision is based on the nearest neighbor principle,

$$Bestmatch = \arg \max_i (d_i(v)) \quad (5)$$

### DYNAMIC TRACKING ALGORITHM



**Figure 9: Definition of occupant in position (top image), occupant out of position (middle image) and occupant in critical out of position (bottom image)**

The fourth and last step covers the recognition and tracking of the occupant's head position in relation to the dashboard surface. For this purpose, an edge detection and a

morphological boarder separation are first carried out for the object of interest. From these results, the shapes of interest (ellipses comparable to a human head) are selected and finally a decision is taken, which of the ellipses detected are in accordance with a human head (and not with similar shapes such as a headrest or a football). The selected shape is then transferred into a Cartesian coordinate system. This data then permits the read out of the actual distance between the head and the place of airbag deployment in an x-, y-, and z-axis (and also to track the head position over a selected period of time).

With a 100Hz system refresh rate of the respective algorithm loop, the occupant's head position is determined and matched into one of three areas: 'in position', 'out of position' and 'critically out of position' (Figure 9). Following completion of this fourth step, all required data is available to take the right decision on how to deploy the airbag (either not at all, with reduced energy, or fully) in line with the Dynamic Automatic Suppression Strategy.

### SYSTEM PERFORMANCE EVALUATION

To evaluate the performance of the optical occupant classification system, as developed by IEE, static classification tests were carried out in-house. For this purpose, a verification of the system according the FMVSS 208 requirements was performed. Subsequently the tests were expanded to include a 'misuse test scenario', as developed by IEE. Tests with separate alternating learning and testing sequences were conducted with an empty seat, both RFIS and FFIS, 'boosters', which are used to give older children a higher sitting position, and with five different population types of humans ranging from the 3 year-old child to the 95-percentile man (Figure 10).



**Figure 10: Overview of different occupant types used for static classification**

To check the reliability of the test system, a range of different environmental influences were applied (i.e. temperature, vibration, contamination of air and camera lens, reflections and scattered light from different sources) as well as various occupant scenarios (i.e. blankets, reflecting glasses, magazines etc.). On top of that a large variety of

torso positions and inclinations of the backrest was compared for the adult occupants.

The results are highly convincing, both for the test series where separate frames were analyzed, and for those with sequences of up to 50 frames (corresponding to a duration of half a second), where a simple filter was applied, significantly improving the results. For the separate frames series, the rate of correctly detected scenarios varies from 99.9% for the FFIS to 92.5% for the adult dummy, and for the sequences this rate varies between 100% (empty seat and FFIS) and 97.8% (adult dummy). The uncertainties in the distinction of persons versus RFIS result from very far forward bending persons, as no history buffer and filtering logic was applied.

% True Class	Estimated Class			
	Empty seat	RFIS	FFIS	P
Empty	97.6	0	0	2.4
RFIS	0	97.9	0	2.1
FFIS	0	0.1	99.9	0
P	0	7.1	0.4	92.5

**Figure 11: Summary of classifier performance based on single images (no history), misuse scenarios included**

% True Class	Estimated Class			
	Empty seat	RFIS	FFIS	P
Empty	100	0	0	0
RFIS	0	98.2	0	1.8
FFIS	0	0	100	0
P	0	2.2	0	97.8

**Figure 12: Summary of classifier performance based on 50 consecutive images, misuse scenarios included (simple filter, no history)**

It is to be expected that filtering strategies based on history buffers will of course eliminate misclassification of adults into the child seat category, as false-true criteria will back up the decision robustness of the system

Further tests show the limits of the test procedure, using living persons as test subject:

Distinction between adjacent size classes (e.g. 5-percentile vs. 50-percentile) is possible at a rate of about 90%.

Distinction between 5-percentile and 95-percentile is possible with almost 100% reliability.

A distinction between six year-old children and small adults is difficult to achieve with high confidence, as the normal distribution of the two classes overlap.

Children on a booster are particularly difficult to determine as their stature is close to the one of the 5%tile female.

% True Class	Estimated Class	
	3 - 6 year or smaller	5%tile or larger
3-6 year on booster	75.6	24.4
3-6 year	90.9	9.1
5%tile	7.5	92.5
50%tile	3.6	96.4
95%tile	0.1	99

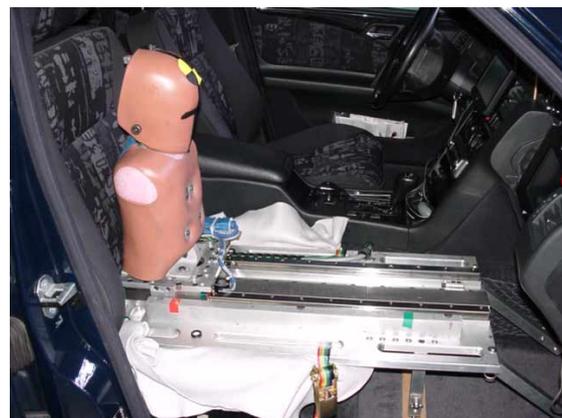
**Figure 13: Summary of classifier performance for different population percentiles**

Beyond occupant classification, as described above, the IEE 3D-Camera can also be used for the head-tracker-test, as separate investigations have shown. Tests had been conducted according to the proposed FMVSS 208 S28.4, DASS test procedure (petition submitted in November 04). For this, a working group called “Smart Vision” (TRW, Siemens VDO, Bosch and IEE) had developed a dynamic OOP test tool to certify the performance of dynamic occupant detection systems in vehicles. Three different analyses – vehicle braking tests, sled tests with braking action, and MADYMO modeling – were conducted to determine the appropriate motion for the DASS tester.

Test results show that

- there is a certain vertical movement of the head, but its vertical position does not change significantly during the tests, and

- the maximum average occupant acceleration relative to vehicle interior is around 4.1 m/s<sup>2</sup>. This determines (an additional safety factor included) a resulting acceleration of the tool of around 4.1 m/s<sup>2</sup> in the specific test setup.



**Figure 14: DASS test tool**

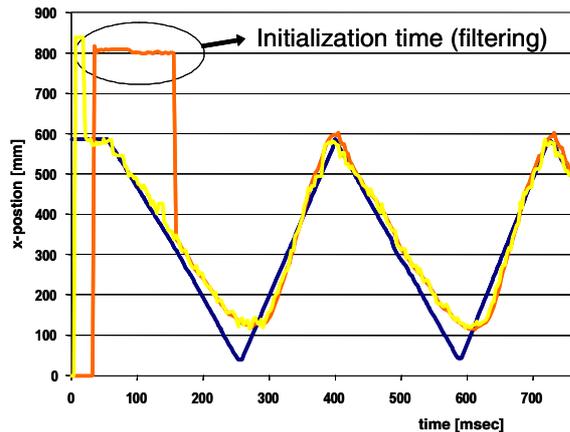
This led to the definition of the following parameters for the DASS Head Tracker Test:

- Linear motion
- Acceleration: 0 to 1.2 g
- Deceleration: 0 to 3 g

- Velocity: 0.5 to 3.1 m/s
- Dummy height: 546 to 635 mm (adjustable)
- Maximum travel: 525 mm

Figure 15 shows a comparison between the positions of

- the test tool,
- the dummy head as recorded by the 3D MLI system, and
- the dummy head as detected by the Head Tracker software.



**Figure 15: Head tracker performance. Motion of test device (blue); Head position seen by the camera (yellow) vs position of head defined by tracker (orange)**

The short initial period of only a few hundred milliseconds, when the traces of tracker and camera deviate, marks the time required by the tracker to verify the correspondence of the identified 'ellipse' and the real object of interest, the head. The virtually perfect coincidence of both traces after this period proves that an optical sensor system, such as the IEE 3D camera, is also applicable for high speed tracking of a moving dummy.

## SUMMARY

As the investigations described here prove, the 3D system developed by IEE provides distance data, which allows a highly precise recognition of the position of an object / a passenger in the FOV (field of view) and thus allows the application of LRD and DASS strategies. 3D data is directly available at the output of the sensor, therefore no additional image processing is required.

The test results also show that vision-based sensors will have their place in the automotive passive safety. Camera systems will be used in future in various passive and active safety applications. Stand-alone camera solutions, as well as a combination of different sensing technologies, will be part of the next generation safety strategies.

## REFERENCES

- [1] J. Krumm and G. Kirk. *Video occupant detection for airbag deployment*. In Proc. IEEE Workshop on Applications of Computer Vision, Oct. 19-21, 1998
- [2] S. Lacroix S. Gautama and M. Devy. *Evaluation of Stereo Matching algorithms for Occupant detection*. In 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.
- [3] P. Faber. *Seat Occupant Detection Inside Vehicles*. In Proc. 4th IEEE Southwest Symposium on Image Analysis and Interpretation, Apr. 2-4, 2000
- [4] P. Faber and W. Förstner. *A System Architecture for an Intelligent Airbag Deployment*. In Proceedings of the IEEE Intelligent Vehicles Symposium 2000, Dearborn (MI), USA, October 3-5, 2000
- [5] R. Lange et al. *Demodulation Pixels in CCD and CMOS Technologies for Time-of-Flight Ranging*. In SPIE Conference on Sensors, Cameras and Systems for Scientific and Industrial Applications, San Jos, USA, 24.-25.1.2001.
- [6] K. Z. Mao. *Orthogonal forward Selection and Backward Elimination Algorithms for Feature Subset Selection*. In IEEE Trans. On Systems, Man, And Cybernetics-Part B: Cybernetics., vol. 34, no.1, Feb. 2004.
- [7] Jürgen Schurmann. *Pattern Classification: Statistical and Neural Network based Approach*, John Wiley and Sons, Inc., New York, 1990.