Robust Occupancy Detection from Stereo Images

Bram Alefs, Markus Clabian, Horst Bischof, Walter Kropatsch and Farid Khairallah

Abstract—**Vehicle occupants that are out-of-position can be deadly injured by the deployment of the air bag in a crash situation. In recent years many different sensors and systems have been proposed to detect the type of occupant and the position of the occupant's head. This paper presents a method for classification and occupant's head detection based on passive stereo vision. The proposed system uses depth surface analysis and scene statistics together with support vector machines for classification and selection of head candidates. Evaluation of the method shows 99% correct for classification and 98% correct for head detection, using large sets of image data, and image sequences recorded in a driving vehicle.**

I. Introduction

AIR bag deployment may cause deadly accidents or severe injuries for certain types of occupants or if the occupant is close to the air bag compartment door. Endangered types of occupants include infants in rear facing infant seats, children and small adults. Moreover any occupant can be killed by air bag deployment if he/she is in an endangered position. The National Highway Traffic Safety Administration (NHTSA) specifies different classes for the occupancy, and out-of-position zones for the human occupants, on which the air bag deployment has to be controlled, see [11] and [13] for a more detailed description.

Detection of the type and position of the vehicle's occupant has been subject to intensive research for the past few years, for example see [2], [3], [6]-[8] and [10]. In order to fulfill future safety requirements, systems are proposed that (1) determine the occupant type, for a fixed set of classes, including adult, child, child seat, empty, and (2) determines the occupant position within the scene. Reference [2], [3] and [6] use stereo vision for classification; [4] and [8] use stereo vision for head detection. Recent studies [10] and [11] include determination of the occupant's posture. Algorithms used in previous studies are applied either directly to the intensity image or the range map. Features include range data, edges, motion [7] and eigenimages [6]. Reference [4], [7] and [8] use support vector machines for occupant classification.

This paper presents a method for classification and real-time head detection that copes with varying illumination, moving shadows and changing contrast, based on a simple stereo camera system. Previously presented methods do not show evaluation results for more complex illumination conditions, such as occur in a driving vehicle. Their robustness is expected to degrade, if additional illumination sources fall out, or if the scene is flooded with direct sunlight. The method presented in this paper only depends on artificial illumination sources, if no natural light is available. It deals with the variety of occupants and the complexity of the illumination condition, using scene statistics. Its robustness is evaluated for large sets of image data recorded in a driving vehicle.

The paper consists of four sections. Section II discusses the method for occupancy classification and head detection. Section III discusses the evaluation results and section IV gives the conclusions.

II. Method

The proposed method consists of two parts: occupancy classification and head detection. Occupancy classification provides the class of the occupant: adult, child, child seat (including rear facing infant seat or RFIS) and empty. Head detection provides the position of the occupant's head within the vehicle.

A. System Setup

The method is designed for a stereo camera system that is mounted in the central upper console on the frontal passenger's side. The camera system has a baseline of a few centimeters and wide angle lenses so that the system observes the entire space at the vehicle's passenger side. The camera system is mounted so that the occupant is unlikely to be occluded by obstacles such as an opened newspaper. Fig. 1 gives an example of images recorded in the vehicle, for different types of occupancy and different illumination conditions. The system includes a passive

Manuscript received March 16, 2004.

Bram Alefs is with Advanced Computer Vision GmbH - ACV, Vienna, Donau-City-Strasse 1; A-1220 Vienna, Austria (http:// www.acv.ac.at; phone: +43-1-269-6255-121; fax: +43-1-269-6255-200; e-mail: bram.alefs@acv.ac.at).

Markus Clabian is also with Advanced Computer Vision GmbH - ACV (address as above, e-mail: markus.clabian@acv.ac.at).

Horst Bischof is with the Institute for Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16 2. OG, A-8010 Graz, Austria (http://www.icg.tu-graz.ac.at ; e-mail: bischof@icg.tu-graz.ac.at).

Walter Kropatsch is with the Pattern Recognition and Image Processing Group, Institute of Computer Aided Automation, Computer Science Department, Vienna University of Technology, Favoritenstrasse 9/1832, 4th floor, A-1040 Vienna, Austria (http://www.prip.tuwien.ac.at; e-mail: krw@prip.tuwien.ac.at).

Farid Khairallah is with TRW Automotive, 24175 Research Drive, Farmington Hills, MI 48335 USA, (e-mail: farid.khairallah@trw.com).

infrared illumination source, for dark outdoor environments.

The camera system provides two intensity images at 30 frames per second. A fast stereo matching algorithm, as discussed in [5] provides the disparity image between the left and right intensity images. Each pixel of the disparity image consists of a value for the distance between corresponding points of the left and right intensity image. The pixel is undefined if no correspondence is found. The range map is estimated from the disparity image, using the calibration parameters.

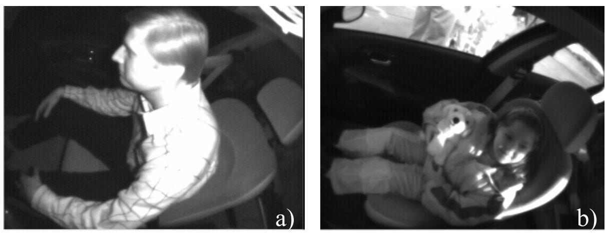

Fig. 1. Example images of different classes: a) adult, b) child. The panels show the left intensity image; the right intensity image looks very similar. The panels show different illumination conditions: indoor (a) and outdoor (b).

B. Classification

The occupancy type is classified with support vector machines (SVM), as they are discussed in [12]. The classification is evaluated for the classes: adult, child, child seat and empty. The class ***adult*** includes human adult occupants of different heights, ranging from 60 to 80 inch. The class ***child*** includes human occupants ranging from 40 to 60 inch height and dummies for children of three and six years old. The class ***child seat*** includes different models and configuration of child seats and infant seats for which the air bag should not deploy, such as the RFIS, occupied and not occupied. The class ***empty*** includes all configurations in which the seat is not occupied.

The input of the classifier consists of features of the entire image. In order to obtain maximal robustness, features are used from the left intensity image and from the disparity image. The accuracy of the classifier and training and evaluation complexity is evaluated for different types of features at different vector lengths. High dimensional feature vectors have a high accuracy, since they contain much information. Training and evaluation time can be extensive, and the system can be overfitted easily. Small feature vectors are evaluated fast and they are not endangered to be overfitted, since the training data is large with respect to the number of parameters of the system. The generalizing property of the classifier is presented by the number of support vectors. The classifier generalizes well, if it has a low number of support vectors.

1) Feature extraction: Since the feature that optimally presents the information is not known, different feature extraction techniques are evaluated. The image is divided in quadratic regions, ranging from 6x4 to 32x24 window positions, so that each region provides one value of the feature vector. Feature extraction techniques include downsampling, texture estimation, gradient estimation and motion. Downsampling is done by nearest neighbor approximation and median estimation. Texture is estimated from the local image variance. Motion is estimated by frame differencing. Extraction of disparity features is more complex, since the disparity image shows undefined regions. Disparity features are extracted with nearest neighbor downsampling, downsampling after a flood-fill operation, downsampling using the median value and motion estimation.

2) Classifier settings: A support vector classifier is designed with a linear and radial basis kernel. The classification system uses a one-to-one voting strategy. For the set of K=4 classes, $N=K(K-1)/2=6$ classifiers are evaluated, and the winner is voted from the results of each one-to-one classifier. Results for the one-to-one classifiers are compared to that of a single perceptron.

C. Head Detection

The method for head detection is designed for occupants of the class adult. A method for the class child can be designed similarly. The method consists of three steps, including generation, selection and tracking of head candidates. The generation step aims to provide at least one candidate for the position of the head. It uses two methods, of which one is based on the depth surface, and the other is based on grouping of edge features in depth slices. The selection step aims to reduce the set of head candidates, so that only the true head candidate is left. It uses scene statistics for large sets of image data. Tracking aims to reduce the number of misses, by assuming that the true head position changes continuously between succeeding frames.

D. Generation of Head Candidates

Generation of head candidates aims to provide at least one candidate at the true head position. It uses two methods, depending on the quality of the disparity image. For disparity images with few undefined regions, head candidates are generated using local maxima in the range map. If the range map is estimated correctly, the head forms a local maximum since it is nearer to the camera than surrounding objects. For disparity images with many undefined regions, a method is proposed that uses grouping of edge features with 3D information, as discussed in [1].

1) Surface-based generation of head candidates consists of three steps: estimation of the range map, determination of the local maxima and analysis of the surrounding region. The range map consists of 25x23 depth values for different image positions. If the disparity is undefined within the region, the range value is approximated from neighboring values or it is set to the background for regions at the rim of the image. In order to speed up, local maxima are estimated with a 5x5 window. For each local maximum, the support region is estimated as the region within a cylinder of 20cm radius, centered at the local maximum, and 20cm depth, with respect to the local maximum. If the area of the region exceeds 150cm^2, its center of mass is denoted as head candidate. Fig. 2 shows an example of surface based generation of head candidates.

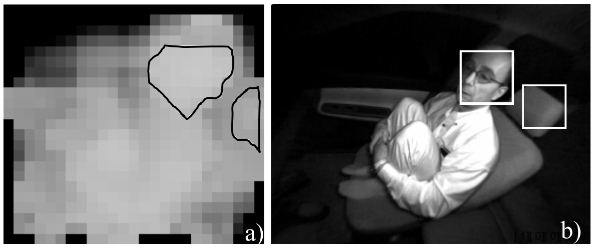

Fig. 2. Surface-based generation of head candidates. a): Range map of 25x23 values; a bright color indicates a position near to the camera. Two support regions are indicated with black polygons. b): Left intensity image of the same cycle. The image regions of the polygons are indicated with white boxes, scaled to 20x20cm.

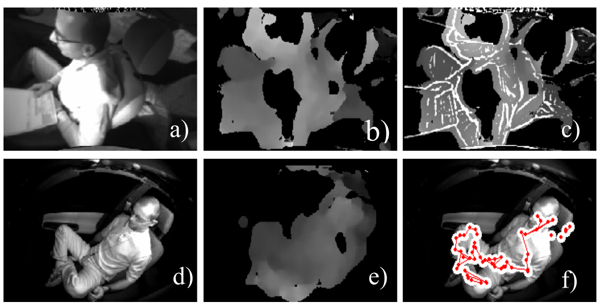

Fig. 3. Edge-based generation of head candidates. Upper row, from left to right: a) Left intensity image; b) Disparity image; c) Superposition of disparity image and edge image. The disparity image is brighter for positions near to the camera. Undefined regions are indicated with black. Lower row: d), e) as a), b); f) feature groups are indicated with connected dots.

2) Edge-based generation of head candidates is introduced for cases in which the head does not form a local maximum in the range map. This may happen if the range map is not properly estimated or if the head is far from the camera. The range map is not properly estimated if the disparity image is insufficiently defined in the head region, due to the lack of texture and to bad illumination conditions. The head is far from the camera for unusual sitting or laying postures if the backrest is reclined. For such positions, the disparity in the head region is small and the local maximum may not be resolved. Edge-based generation of head candidates copes with these situations by regarding edge features rather than the local maximum. Edge features are extracted from the intensity image, using a LOG-filter, and they are grouped in 3D space. The feature density is estimated for cuboids of 40x40x20cm in width, height and depth, and regions with a high feature density are approximated with ellipsoids. Objects are formed by connecting ellipsoids, so that the uppermost ellipsoid of each object provides one head candidate. Fig. 3 shows an example of edge and disparity features. Fig. 3f) indicates the center positions of the connected ellipsoids.

E. *Selection of Head Candidates*

Selection of head candidates aims to provide the true head candidate from a set of candidates, by rejection of false candidates. Selection criteria are estimated using scene statistics for a set of 4000 representative images. The dataset is discussed in section III.C. For each image, a set of candidates is generated using the surface-based generation method and the true head candidate is selected manually. Based on scene statistics of the true head candidate, false candidates are rejected. Selection criteria are based on the absolute position of the head in 3D space, the relative position between candidates and the surface geometry.

1) The 3D head position is estimated for all true head candidates of the data set. The probability function is derived from the histogram of the support region of the true head candidates, using bins spread over the 3D space. In order to speed up the evaluation, bins are positioned equally over the image coordinates, for 25x23 image positions and for 18 depth steps, ranging from 35cm to 120cm. The histogram is smoothed and manually adjusted to exclude positions of false candidates that origin from the B-console, the leg region and the sun visor. A probability value is derived for each of the 25x23x18 bins, so that 22% have a value larger than 0. All candidates with zero probability are rejected. Remaining candidates include the occupant's head, parts of the body, the head rest and other objects within the compartment.

2) Selection based on relative position is applied for the remaining candidates, which typically include the head and the head rest. In order to eliminate the candidate of the head rest, the center of mass of the candidates on the right side of the image is estimated. A selection plane is designed, so that candidates right and behind from the center of mass are eliminated.

3) The surface geometry is used to select the remaining candidates. Statistical knowledge is obtained about the surface geometry of the true head candidate. False candidates with dissimilar surface geometry are rejected. The surface geometry includes values for area, shape and curvature. The area is estimated by summation of the area for each element of the support region. The shape is estimated by singular value decomposition for data represented in the plane of the image. The curvature is estimated with a second order polynomial approximation of the depth as function of the image coordinates. Since most false candidates show similar statistics, candidates are rejected only if one of the parameters is out of range.

F. *Tracking*

Tracking is introduced in order to reduce the number of misses. The true head candidate is missed, if none of the head candidates shows the true head position, or no candidate is left at all. Typically, misses occur if the head is invisible due to bad illumination conditions or it is occluded by objects, arms, legs etc. A 3D point tracker as discussed in [9] is designed to cope with temporal invisibility and occlusions. Head candidates are tracked by using the correspondence between the candidates of succeeding frames, while counting the observation history for each of the candidates. If a candidate disappears, additional candidates are predicted at the previous position. The duration of the prediction depends on the number of previous observations, which is limited to N_{max}. Correspondences are searched within 20cm, i.e. the head moves with maximally 6m/s, for a frame rate of 30fps. Motion models are not included, since most changes are caused by inaccurate positioning rather than by the head movement itself.

III. Results

A. System Setup

Results are obtained with three different camera systems, of which two are mounted in a vehicle, and one is mounted in the office. None of the camera systems provided color information. Table 1 gives an overview. The illumination conditions are specified as ***indoor***, if the infrared illumination source is dominant or ***outdoor*** if (direct) sunlight is dominant (see fig. 1). The illumination condition in the office was simulated by changing the exposure control, so that images range from entirely dark to entirely bright. For all camera setups and all sitting postures and seat configurations, the occupant was at least partly visible, and the head was within the field of view. Occlusions were mainly at the body of the occupant. The head was occluded if hands were put between the head and the camera, or for unusual sitting postures, such as for frontal sitting occupants that grasp something from the back bench. Left and right intensity images were captured at 30fps, at a resolution of 320x240 pixels. The disparity image was provided at the same frame rate, so that maximally 250x230 pixels at center of the image are defined.

B. Classification

The classifier was trained and evaluated with two different strategies. The first strategy evaluates the classification accuracy using the leave-one-out technique, so that the system is repeatedly trained on a randomly chosen set of samples and it is evaluated for the remainder. The second strategy evaluates the robustness of the system by leaving out specific groups of training data. This strategy helps to understand the sensitivity for certain types of training data and it gives insight in the generalization of the classifier.

Image data was recorded with the camera system mounted in Vehicle I, as shown in Table I. Image data includes recordings of 34 adult occupants; 49 children and 2 child dummies; 9 child seat types and an empty seat in various configurations. For all classes, the illumination condition was indoor (76%) and outdoor (24%). Human occupants were recorded continuously, while the vehicle was driving, or the occupant moved in forward and backward direction. Recordings of adults showed risen arms, out-of-positions and an opened newspaper. Recordings of children showed standing and laying postures. They include objects such as cuddly toys. Static occupancies, such as dummies or child seats and empty were recorded in discrete steps, while the seat position or sitting configuration was (automatically) changed. Continuous recordings were sampled to one feature vector per ten frames. Discrete recordings were sampled to one feature vector per frame. Totally, 3228 samples were extracted, of which 46% of the class adult, 20% of the class child, 24% of the class child seat and 10% of the class empty.

TABLE I
CAMERA SYSTEMS & MOUNTING

Mounting	Sensors	Baseline (mm)	Focal Length (mm)
Vehicle I	CMOS, low contrast	27	3.0
Vehicle II	CMOS, high contrast	35	2.2
Office	CCD	54	2.6

The field of view depends on the focal length and the chip size. Side regions were removed in case of large lens distortions. The effective field of view is largest for the camera system used in Vehicle II.

TABLE II
FEATURE EXTRACTION

	Intensity Image %Correct (# Support Vectors)	Disparity Image %Correct (# Support Vectors)
Downsampling		
Nearest Neighbor	95.1 (359)	94.0 (400)
Flood-fill	-	96.8 (319)
Median	95.8 (331)	97.0 (317)
Texture	95.4 (342)	-
Gradient	95.5 (342)	-
Motion	66.0 (734)	82.3 (985)

Results for 3228 samples with different feature extraction techniques, for a vector length of 192. Downsampling after flood-fill is applied to the disparity image only. Texture and gradient are only applied to the intensity image only. The percentage correct indicates the average results for all classes, the number of support vectors include support vectors for all classifiers.

1) Leave-one-out. The system is repeatedly trained on 50% or 99% of the data, and evaluated for the remainder, so that statistically each sample is evaluated at least once. Classification results were obtained with the support vector machine, with a linear kernel. One-to-one classification with a perceptron showed similar or slightly worse results, depending on the training strategy. Table II shows results for different features. Best performance are obtained with downsampling, using median gray and disparity values for 16x12 window positions and for all frames of the sample interval. Shorter feature vectors resulted in an increase of the error rate and an increase in the number of support vectors. Longer feature vectors showed a similar error rate and an increased number of support vectors. Table III shows the confusion matrix for a combined vector, using downsampling to 16x12 median disparity values and 12x9 median intensity values. The average result for this combination is 99.1% correct.

2) Robustness. The robustness was evaluated for variation of illumination, variation of individuals and variation of positions and occlusions. Specific groups of data were left out of training, so that eventually the system breaks down due to the lack of training data. Results were obtained with a reduced set of training data, for which the system starts to degrade. Further reduction led to a full degradation of the system. Totally, 77 recordings were specified with different individuals, illumination conditions and child seat models. The system was slightly degraded, if one of the five recordings of class adult with outdoor illumination was left out. The system was largely degraded if one of the recordings of class child seat and empty with outdoor illumination was left out. Results were not degraded if one of the recordings with indoor illumination was left out, for any class. Table IV shows results if all individuals of the height group were left out of training. Results degrade for the group of tallest adults and children only.

TABLE III
CLASSIFICATION RESULTS

	Adult	Child	Child Seat	Empty	# Samples
Adult	99.2	0.0	0.7	0.1	1503
Child	0.2	99.6	0.0	0.2	650
Child Seat	1.4	0.0	98.6	0.0	764
Empty	0.7	0.3	0.0	99.0	311

Confusion matrix for 3228 samples, using a combined feature vector of 16x12 median disparity values and 12x9 median intensity values. The rows indicate the percentage of samples that are classified and the total number of samples for each class. The columns indicate the output of the classifier. The total number of support vectors is 407.

TABLE IV
ROBUSTNESS OF HEIGHT GROUPS

Child Group (height)	%Correct (#Individuals)	Adult Group (height)	%Correct (# Individuals)
40-44 inch	100% (2)	60-64 inch	97 % (11)
45-49 inch	100% (5)	65-69 inch	96 % (8)
50-54 inch	99% (18)	70-74 inch	97 % (13)
55-59 inch	84% (19)	>74 inch	81 % (2)

Results for different height groups of class adult and class child. Each result indicate the percentage correct for the samples of each group, if the entire group is left out of training. Occupants of child group 55-59 inch are confused with adults, if left out of training.

TABLE V
GENERATION OF HEAD CANDIDATES

Method	True Head Candidate (%Images)	Number of Candidates (Average)	Distance to Ground Truth (pixels)
Surface-based	64.1	11.5	33
Edge-based I	83.6	1.9	17
Edge-based II	93.8	4.6	16

Results for 1156 images with simulated illumination conditions, in which the head was visible for a human observer.

C. *Head Detection*

Head detection is evaluated for different sets of data, recorded with the camera system mounted in Vehicle II and in the office. The image data of Vehicle II consists of still images with a large variety between occupants, position and sitting postures, and sequences. In order to evaluate bad illumination conditions, image data is recorded in the office so that illumination is simulated by changing the exposure control. Office recordings are used to compare results for the generation of head candidates. The selection of head candidates is evaluated for the discrete image data used for training and for continuous recordings in a driving vehicle. Tracking is evaluated for continuous recordings only.

1) Generation of head candidates depends on the illumination condition and the occupant's posture. The dependence on the illumination condition is evaluated for data recorded in the office. The dependence on the occupant's posture is evaluated for data recorded in Vehicle II. Table V shows the results for 1156 images recorded in the office, for which the head was visible for a human observer. Images for which the human observer could not indicate the head position because of bad illumination were left out. The method Edge-based I uses cuboids of 40x40x20cm, as discussed above; the method Edge-based II uses cuboids of 10x10x5cm, so that smaller and partly visible objects are included. The edge based methods perform better for bad illumination conditions. For data recorded in Vehicle II, edge-based generation of head candidates performs better for unusual sitting postures and bad illumination conditions. For the majority of images the performance is similar, and the surface-based method is used on from here, because it is faster. Misses occur if the head is far from the camera, or if the head region is badly illuminated.

2) Selection of head candidates is evaluated using data recorded in Vehicle II. The data consist of 4000 still images of 21 adult occupants, in different sitting postures, seat configurations and illumination conditions: 200 images show a dummy occupant in various positions; 200 images show human occupants, reading a newspaper; 400 images show human occupants with a bad illumination condition or in an unusual sitting posture; 400 images show human occupants with risen arms, opening the sun visor; 400 images show occupants leaning backward, with legs on the dashboard; 2400 images show human occupants in normal sitting posture, with out-of-positions.

TABLE VI
SELECTION OF HEAD CANDIDATES

Method	Unique Detection (%)	Ambiguous Detection (%) (Number of Candidates)	Missed (%)
Generation	0	94 (8.8)	6
Selection I	40	52 (2.4)	8
Selection II	73	17 (2.3)	10
Selection III	76	13 (2.2)	11

Results for 4000 still images. The head is uniquely detected if a single candidate is found that coincide with the true head position. It is ambiguously detected if one of the candidates provides the true head position and it is missed, if none of the candidates provide the true head position, or if no candidate is found at all.

TABLE VII
TRACKING

Method	Unique Detection (%)	Ambiguous Detection (%)	Missed (%)
Selection	96.7	0.2	3.1
Tracking I	93.7	3.6	2.7
Tracking II	96.6	1.4	2.0
Tracking III	97.8	0.6	1.6

Results for 12 sequences of 30s. See Table VI for a description of the columns. The selection method coincides with Selection II of Table VI, now evaluated for the continuous recordings.

Table VI shows results for different selection steps. The detection result is either 'unique', 'ambiguous' or 'missed'. Generation results are obtained with the surface-based method. Selection I includes selection of the 3D head position only; Selection II includes additional selection of the relative position and Selection III includes selection of the surface geometry also. The generation method yields no unique detections. Misses are found due to bad illumination and unusual postures, as discussed above. The number of candidates is strongly reduced after Selection I, and on from Selection II, the head is uniquely detected for the majority of images. Results for Selection II for continuous recordings are shown in Table VII.

3) Tracking is evaluated for 12 sequences of 30 seconds with adult occupants in a driving vehicle, using the camera system mounted in Vehicle II. The illumination condition was outdoor; no additional infrared illumination source is used. Images show a moving background, direct sunlight and moving shadow patterns. The occupant was stationary, or moved towards out-of-position. Table VII shows results after Selection II, as discussed above, and for different values for the maximal observation history N_{max}: 1 frame (33ms) for Tracking I, 10 frames (0.33s) for Tracking II and 100 frames (3.3s) for Tracking III. Selection without tracking shows less misses as for data shown in Table VI, since image data does not include unusual sitting postures. The percentage of misses is reduced with a factor 2, on from a typical duration of the occlusion or illumination condition of one second. Fig. 4 shows examples of head detection in a driving vehicle.

Fig. 4. Examples of head detection in a driving vehicle, including direct sunlight and moving shadow patterns. The elliptical contour is estimated from the 3D-head position, using radii of 12cm and 9cm.

IV. Conclusions

Occupancy detection from stereo images is sensitive for variations between occupants, seat configurations and for different illumination conditions. A method for occupancy classification and head detection is proposed that performs well for large sets of image data, with many occupants, unusual sitting postures and bad illumination conditions. The method makes intensively use of statistical information about the observed scene and the occupants. Occupants and sitting postures are varied and bad illumination conditions are recorded and simulated. Evaluation shows 99% correct classification and 98% correct head detections for data recorded in a driving vehicle, without additional illumination sources.

Acknowledgment

This work has been carried out within the K plus Competence Center ADVANCED COMPUTER VISION. This work was funded from the K plus Program. We thank A. Leonardis and V. Hlavac for their helpful comments.

References

[1] B. G. Alefs, M. Clabian, H. Bischof, W. Kropatsch; Head detection based on feature grouping in depth slices; in Proceedings Workshop of the Austrian Association for Pattern Recognition (ÖAGM/AAPR); pp. 275-280; Austrian Computer Ass. 2003.

[2] M. Devy , A. Giralt , A. Marin-Hernandez ; Detection and classification of passenger seat occupancy using stereovision; in Proceedings IEEE Intelligent Vehicles Symposium; pp. 714-719; 2000.

[3] P. Faber, W. Foerstner; A system architecture for an intelligent airbag deployment; in IEEE Intelligent Vehicle Symposium; pp. 70-75; 2000.

[4] C. Koch, T. J. Ellis, A. Georgiadis; Real-time occupant classificiation in high dynamic range environments; in IEEE inteligent vehicle symposium; vol. IV'2002; 2002.

[5] K. Konolige; Small vision systems: hardware and implementation; in Eighth International Symposium on Robotics Research; 1997. http://www.ai.sri.com/~konolige/svs/

[6] J. Krumm, G. Kirk; Video quantitative evaluation of motion detection algorithms for surveillance applications; in 4th IEEE Workshop on Applications of Computer Vision; pp. 1-21; 1998.

[7] Y. Owechko, N. Srinivasa, S. Medasani, R. Boscolo; Vision-based fusion system for smart airbag applications; in IEEE Intelligent Vehicle Symposium; vol. IV'2002; 2002.

[8] R. Reyna, A. Giralt, D. Esteve; Head detection inside vehicles with a modified SVM for safer airbags; in IEEE Intelligent Transportation Systems Conference Proceedings; pp. 268-272; 2001.

[9] V. Salari, I. K. Sethi; Feature point correspondence in the presence of occlusion; in IEEE Transactions on Pattern Analysis and Machine Intelligence; vol. 12; pp. 87-91; 1990.

[10] T. D. Schoenmackers, M. M. Trivedi; Real-time stereo-based vehicle occupant posture determination for intelligent airbag deployment; in IEEE International Conference on Intelligent Vehicles; pp. 570-574; 2003.

[11] M. Trivedi, S. Y. Cheng, C. Childers, S. Krotosky; Occupant posture analysis with stereo and thermal infrared video; in IEEE Transactions Vehicular Technology, Special Issue on In Vehicle Computer Vision Systems, 2003, submitted for publication.

[12] B. Schölkopf, C. J. C. Burges, A. J. Smola; *Advances in kernel methods*, support vector learning; Cambridge, Massachusetts: MIT press, 1998, ch. 1.

[13] Federal Motor Vehicle Safety Standards; Occupant crash protection; final rule; Department of Transportation; Federal Register; vol. 65; no. 93; pp. 30680-30770; 2000.