An Occupant Classification System - Eigen Shapes or Knowledge-Based Features

S. Burak Goktuk[1,2], Abbas Rafii[1]
[1]*Canesta Inc.*, [2]*Ojos Inc.*
burak@ojos-inc.com, arafii@canesta.com

Abstract

This paper describes an occupant classification system based on Eigen shapes and support vector machines classification using 3D data. The occupant classification system is used to classify an occupant type into an adult, a child, a child seat, or an object. The inputs are depth images from a time-of-flight based depth camera. The images are first segmented from the background and normalized for three-dimensional translations. The segmented images are projected into Eigen shapes that are constructed from a training set. The projections are used as feature vectors for a support vector machines classification algorithm. The Eigen features and support vector machines complement each other since the former is a linear transformation to reduce the dimensionality of the space, while the latter can deduce the non-linear aspects of these lower-dimensional features. Our experiments lead to several conclusions. First, we compare between Eigen shapes and knowledge based features, e.g. computer generated versus human generated features. The Eigen shapes provides better results compared to various combinations of knowledge-based features. The combination of Eigen shapes and knowledge-based features provide the best results, since both of these features capture different characteristics of images. The Comparison with intensity based analysis shows that the depth images are more suitable for this application. The support vector machines have also been shown to be superior to several other classification algorithms. The system provides more than 98 percent recognition rate with depth images as input. The failure cases include extreme deformations of the occupant that were not part of the training set.

1. Introduction

This paper describes a time-of-flight based sensor and system for occupant classification in intelligent airbags. Although airbags save many thousands of lives every year, they are also the cause of death or injury when a child is sitting in front, or when the occupant's head is close to the airbag. As a consequence, the U.S. Government and others have enacted the Federal Motor Vehicle Safety regulations ("FMVSS 208") that require automakers to introduce advanced airbag systems that deploy based on the size and position of the occupant at the time of the crash.

Occupant monitoring is a challenging problem in many ways. First, the illumination conditions vary enormously, from no light condition at night to very bright sunlight conditions. The accuracy of the system is very important since its results relate to the safety of occupants. On the practical side, the system needs to be low-cost, have small form factor and function inconspicuously. In addition, the system should be able to deal with many challenging cases (Figure 1) in real-time.

These challenges impose various requirements on a vision-based system. First, the system should have its own light source in order to work at night. Next, the result should be invariant to changes in ambient light condition. Good accuracy and robustness of the system are also necessary for detecting the position and type of the occupant irrespective of surface and color conditions. The above requirements rule out 2D camera based system, since it depends on the lighting conditions of the environment and the color and texture of the occupant's skin and clothing. The 3D sensors, on the other hand, are very attractive since they provide the range data, which is independent of color of objects, and the ambient conditions in the scene. More specifically, time-of-flight based 3D sensor is preferable since it works reliably on textured and non-textured surfaces, it works regardless of the ambient conditions, and it can be packaged in a small form factor since it does not require a baseline.

Two tasks are necessary to monitor an occupant. First, the occupant's type has to be classified as adult, small adult, child, unspecified object or empty seat, and next if the occupant is a person, the occupant's head position needs to be tracked. We have provided a practical and robust solution to the head tracking problem using a recognition and interpolation framework in [1]. In this paper, we provide a solution to the occupant classification problem.

The flow of the system is given in Figure 3. The inputs to the occupant classification system are depth images from the time-of-flight depth sensor. The depth images are first segmented and normalized. Then, Principal Component Analysis (PCA) is applied on the normalized depth images to obtain the occupant Eigen shapes. The projection into the Eigen shape is used as a feature vector on each training and testing sample. A Support Vector Machine (SVM) is applied on the training set for each of the classes by a one-to-many scheme. Finally, the classification is applied on the testing dataset, and a label is assigned to each test image.

Recently, research has been conducted in the area of occupant classification. The current products are all weight sensor based [2,3]. This technique does not perform well in many cases. For instance, the pressure that a heavy child seat provides might be higher than a small-sized adult that is leaning forward. Vision-based occupant classification research has gotten attention in the last few years. In [4], Farmer and Jain proposed an intensity image based technique. Although they reached recognition rate of 95% at various controlled lighting conditions, such a system can not be used in a practical setting due to the large variations in the lighting conditions in a car. In [5], Kong et al. proposed a technique on stereo cameras. They used combinations of the down-sampled images, edge images or wavelets as features and trained a neural network for classification. They obtained a recognition rate of 70%. We contribute this lower rate mostly due to the fact that they use stereo cameras which would not perform under textureless surfaces. Eigen Shapes and SVM have been used in many computer vision applications [6,7,8,9,10]. The combination of the two and their potential improvement over other methods have not been exploited much. This issue is going to be raised throughout this paper, and we provide experiments that support why they complement each other well.

We have various contributions where we differentiate from previous work. First, we experiment a new type of depth sensor, which is appropriate for an occupant classification system. Second, we provide a full system that involves a new type of depth sensor, image processing and computer vision techniques for normalization via projection between 2D and 3D, pattern processing techniques for feature extraction, and statistical analysis techniques for classification. Third, we provide experiments that optimize the input type, types of features, and classification algorithms to be used in an occupant classification application. Finally, we emphasize that the features and classification algorithm go hand-in-hand for a good performance, and that the research should also focus on good combinations of the two, rather than just optimizing one or the other. It has been shown in literature that Eigen shapes provide inferior results compared to many other feature extraction techniques [11]. However, these works usually use a simplified classifier, such as nearest neighbor classifier. A classifier like SVM finds an optimal surface to differentiate between the feature sets of different classes. The performance of the SVM depends heavily on the dimension of the feature vector. As the dimension of the feature vector increase, the number of samples required to train the system increase exponentially ('Curse of Dimensionality')[12]. By using Eigen shapes, the dimension of the feature space is reduced without compromising much from the content. The lower-dimensional features of Eigen shapes boost the performance of the SVM Classification. As such, the combination of Eigen shapes and SVM is suitable for a classification problem.

In this paper, we analyze various aspects of the framework. First, we apply our methodology to both intensity and depth images, and show that the depth data provides superior results. Second, we compare the classification success of Eigen shapes method to various other feature extraction techniques. Third, we apply various linear and non-linear classification schemes. We show that the Eigen shapes make a better match-up with non-linear classifiers. Next, we show that SVM provide better results compared to other classification schemes such as linear discriminant analysis, nearest neighbor classifier, and n-nearest neighbor classifier.

The paper is structured as follows: First, we give an overview of the working principle of the time-of-flight depth sensor. Next, we give an overview of the occupant classification algorithms. We describe the object segmentation and normalization, depth-based Eigen shapes, and classification algorithm respectively. Next, we describe the experiments and results. Finally we provide our conclusions and discussions.

Figure 1. Examples of challenging cases. The first row shows the brightness images, whereas the second row shows the depth images. The depth images are color-coded such that the color becomes darker as the object is closer.

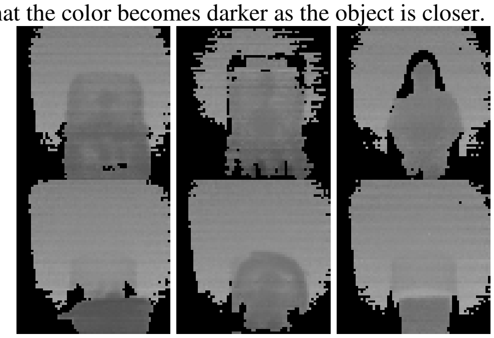

Figure 2. Example Depth images. Top Row images(left to right): A child seat, a child seat with the hands up, a small sized lady. Bottom row: Another type of child seat (rear facing), A front facing child seat, and a box.

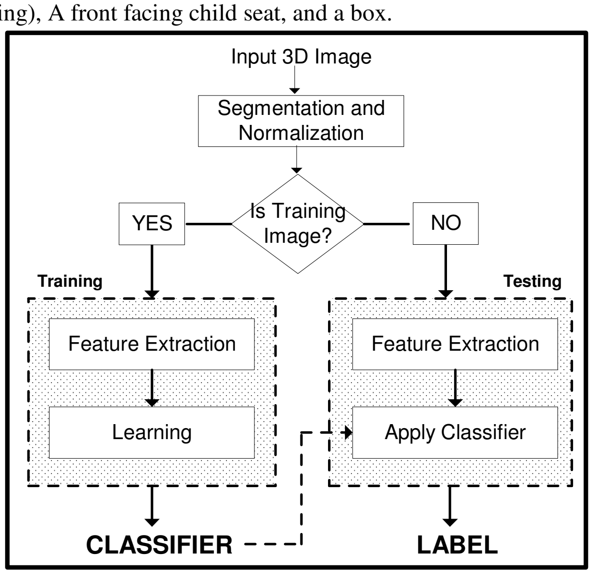

Figure 3. Flow of the Algorithm

2. Time-Of-Flight Depth Sensor

The time-of-flight imaging system was used in our experiments to produce the depth map. The system consists of a modulated light source such as a laser, a CMOS sensor consisting of an array of pixels each capable of detecting the intensity and phase of the incoming light, and an optical system for focusing. Distance is computed from the phase of the modulation envelope of transmitted infrared light as received at each pixel [13].

A depth image is constructed by measuring the distance d at every pixel of a 64x64 pixel array. Similarly, a brightness image is constructed via measuring the amplitude of the received light at every pixel. The phase detection was implemented in CMOS circuitry as described in [13, 14]. The system has a resolution of about 1 cm. [15]

The ambient lighting conditions in the scene would adversely affect the performance of an active system. Our sensor, although an active sensor, has various means to cancel the effects of ambient light in the scene. First, the sensor has a light filter that only passes a band around the wavelength of the active light. Second, the chip involves electronic circuitry that differentiates the modulated component of the light, from its baseline component. Third, we have algorithmic techniques that remove the noise artifacts of ambient light [15]. The combination of the three assures that the system can operate outdoors.

The time of flight sensor is different from other depth sensor in various ways, and more suitable for an occupant classification system. First, the system can work both on day and night, regardless of the lighting conditions in the scene. Second, unlike stereo, it is texture independent. Similarly, the system does not necessitate a baseline between the light source and the camera, and as such there is no parallax shadows. The depth calculation is done in the CMOS circuitry, freeing up CPU time for application level processing. It uses diffused floodlight, as opposed to structured light. This provides an advantage over structured light systems, since there is no moving light part and no eye-safety problem. Finally, the depth sensor is implemented on a CMOS chip, and this provides a small, inexpensive and relatively high-resolution depth sensor for an occupant classification system.

Figure 2 shows example depth images of a person, child seat and objects. Here, we color-code the images such that objects become darker as they are closer to the camera. The output depth images are used as input to our classification algorithm as described next.

3. Overview of the Occupant Classification Algorithm

An Overview of the algorithm is given in Figure 3. The input is the depth image from the time-of-flight sensor. The image is first segmented from the background, and normalized for 3D translations in space. The normalization is accomplished by constructing points in 3D, and by back projecting them on a virtual camera of fixed location.

The system has two modes, a training mode, and a testing mode. In both modes, a feature vector is created from the normalized depth image. We use Eigen shapes and knowledge-based features as described in Section 3.2.

The features of the training dataset are used to learn a Support Vector Machines classifier in the training stage (Section 3.3). The learnt classifier is applied on the features of the test dataset in a one-to-many fashion, and the final label of the test image is determined.

3.1. Depth Image Segmentation and Normalization

A segmented image is created from every image. The segmented image should contain only the objects in the picture. Let I be the depth image obtained from the depth sensor. The segmentation is accomplished by subtracting a background depth image I_B. I_B contains only the seat and interior of the car from the camera's view point. Such an image can be obtained in the factory, or automatically when the car door is opened. Thresholding is applied on the difference image and a segmented image I_S is obtained as follows:

$$I_S(r,c) = \begin{cases} I(r,c) & if\ |I(r,c) - I_B(r,c)| > T \\ 0 & if\ |I(r,c) - I_B(r,c)| < T \end{cases}$$

where $I_S(r,c)$, $I(r,c)$ and $I_B(r,c)$ are the corresponding values of the pixel at row r, and column c in images I_S, I and I_B respectively. T is the threshold used in this operation, and should be chosen well above the noise limit of the sensor. A threshold level equivalent to 10 cm was used in our experiments.

Once the image is segmented, the segmented image I_S is used to obtain a normalized image I_N. First, each non-zero pixel in I_S is projected out to 3D space to generate a 3D point for each pixel (r,c):

$$P(r,c) = \prod_C^{-1}(I_S(r,c))$$

where $P(r,c)$ is a 3 by 1vector denoting the location corresponding to pixel (r,c) in image I_S. $\prod_C$ is the projection matrix from camera world coordinates to the camera C. Observe that obtaining $\prod_C^{-1}$ is trivial since the depth is already provided by the camera. Next, the points are projected back to a virtual camera that is fixed distance from the seat:

$$I_V(r,c) = \prod_V(P(r,c))$$

where V is the virtual camera, $\prod_V$ is the projection matrix from camera world coordinates to the camera V, and I_V is the virtual camera image. The camera C, and the virtual camera V share a common optical axis. Next, the center row (r_c) and the center column (c_c) location of the image I_V is found:

$$r_c = \frac{1}{N}\sum_{(r,c),I(r,c)>0} r \ , \ c_c = \frac{1}{N}\sum_{(r,c),I(r,c)>0} c$$

Finally, the normalized image I_N is obtained by normalizing r_c and c_c to the center of the image:

$$I_N(r,c) = I_V(r + \frac{R}{2} - r_C, c + \frac{C}{2} - c_C)$$

where R and C are number of rows and columns respectively. This ensures that the objects in the image I_N are normalized with respect to 3D translation in X, Y, and Z. A down-sampled version of I_N is also calculated and kept in image I_D.

3.2. Eigen Shape Analysis and Knowledge-Based Features

We calculate and experiment with two types of features. The first one is Eigen shape analysis, where principal component analysis is applied to get the most relevant projections of the data in the least squares sense[16]. The second one is various sets of knowledge-based features, where the heuristics about the application is being used to get the features. We describe both of these features next.

The downsampled, normalized image I_D is used as input to create the Eigen Shapes. First, all pixel depth values of image I_D are listed into the vector v_D. A matrix A is constructed by listing the training vectors v_D's as columns of the matrix. Singular value decomposition is applied on matrix A:

$$A = U.\Lambda.V^T$$

where Λ is a diagonal matrix of the same dimension as A and with nonnegative diagonal elements in decreasing order, and U and V are unitary matrices. The first n columns of U are used as the principal components (U_n), and each vector v_D is projected onto the principal components to create the Eigen shape based feature vector f_E:

$$f_E = U_n^T.v_D$$

The knowledge based features have been generated through our heuristics. We divided the features into different sets, and we provide the performance with combinations of these features in our experiments. The first set contained the following features:

- Total surface area of the object in image I_N.
- Average row and column location of the object in image I_N.
- Statistics about the depth values, such as average, minimum, maximum, and standard deviation of depth values in image I_N.

- Starting and ending row and column location of the object in image I_N.
- A ratio of number of valid columns in a starting row and in a middle row. Observe that this feature gives information regarding the existence of a human since it gives a ratio of his head's width versus his body's width.

The second set contained shape features, such as zero, first and second order moment invariants [17], and a 3x3 downsampled image of I_N. The third set contained a depth histogram of the image I_N. The fourth set contained statistics on the row values of image I_N, such as the start and end location, and average depth value of the valid pixels in a row. We experimented with various combinations of these feature sets (Section 4).

3.3. Support Vector Machines (SVM) Classification for Multiple Classes

The feature vectors of the training data are used to train a Support Vector Machines (SVM) classifier. Each image is labeled as one of 3 categories: Adult, child, and object. Child set includes cases from front and rear facing child seat, as well as child sitting on the car seat. The object set includes cases with boxes, and shopping bags, and empty car seat. Examples of these cases are illustrated in Figure 2. Let us first consider two-class classification problem, i.e. classification between adult set vs. all other sets.

Given the feature vectors for the images, the optimum classifier has to be obtained using the training data set. The goal is to find a separation function that can be deduced from the known data points and generalizes well on the unknown examples. Proposed first by Vapnik [18], SVM classifier aims to find the optimal differentiating hypersurface between the two classes. In the general case, the optimal hypersurface is the one that not only correctly classifies the data, but also maximizes the margin of the closest data points to the hypersurface. The mathematical details of finding this hypersurface can be found in [18]. Let z_f be the signed distance of a feature vector f to this hypersurface, i.e. the sign depends on which side of the hypersurface the feature vector falls in. The sign of z_f on a test case determines the classification of the case in the two-class classification problem.

The theory of SVM needs to be extended to multiple classes for the occupant classification task. In the multiclass problem, best differentiating SVM hypersurface is deduced for each class, where this hypersurface distinguishes that particular class from the rest of the data. This is called one to many scheme. The training is complete once a hypersurface is determined for each class.

While testing a new case, the location of the data is first determined with respect to each hypersurface. Let z_i be the distance of the new data point to the i^{th} class distinguishing hypersurface. The probability that the new data belongs to the i^{th} class is assigned by,

$$P(i) = \frac{e^{Z_i}}{\sum_k e^{Z_k}}$$

The proposed probability assignment method is chosen mainly due to the fact that the exponential function is continuous on positive and negative values of z_i. The class with the highest probability is assigned as the system's final choice among multiple classes.

4. Experiments and Results

We collected a dataset of 12 people, three child dolls, and two child seats, and various boxes and bags. Separate training and testing data sets were obtained using the depth sensor. 4 of the people included in the testing set were not included in the training set. The people were asked to perform various movements during the data capture. Some examples are provided in Figure 1, 2, 4 and 5. The training set included images of 399 adults, 332 children or child seats, and 207 objects. The testing set included images of 1261 adults, 297 children or child seats, and 97 objects.

We divided our experiments to evaluate the following categories: Choice of input type, choice of features, and choice of classification algorithm. In the next sections, we provide our results using different input types, i.e. depth versus intensity images, using different feature sets, i.e. Eigen Shapes versus knowledge-based features, and finally using different classification schemes, i.e. SVM versus others.

4.1. Depth Images versus Intensity Images

The depth sensor provides both intensity and depth images at every frame. This feature gave us a chance to compare the performance of the depth and intensity images obtained from the same content. We obtained Eigen shapes as the features, and applied SVM with a kernel of exponential radial basis functions [19]. We put aside the fact that the intensity images are lighting dependent, and that they would not work at night without additional lighting. Table 1 lists the results of experiments with the intensity images and depth images. The columns illustrate the number of misclassified adult, child, and object classes, and the total error rates.

	Adults	Child	Object	Error (%)
Depth	11	16	0	1.63
Intensity	1	52	4	3.44

Table 1. Misclassified cases with Depth and Intensity Images

	Adults	Child	Object	Error(%)
Eigen	11	16	0	1.63
Set 1	27	48	4	4.77
Set 2	14	70	0	5.07
Set 3	175	76	20	16.36
Set 4	11	42	0	3.20
Set 1+2	36	27	0	3.80
Set 1+2+3	19	20	0	2.36
Set 1+2+4	9	32	0	2.48
Set 1+2+3+4	10	27	0	2.23
Eigen+ Set 1+2+3+4	3	21	0	1.45

Table 2. Misclassified cases with Different Features

	Adults	Child	Object	Error(%)
SVM – Linear	129	31	0	9.66
SVM – RBF (2)	52	12	0	4.47
SVM – RBF (4)	61	17	0	4.71
SVM – ERBF (2)	11	16	0	1.63
SVM – ERBF (4)	21	16	0	3.20
LDA	56	20	4	4.83
1-NN	42	25	0	4.05
8-NN	90	25	0	6.94
15-NN	91	26	0	7.07

Table 3. Misclassified cases with Different Classifiers

Our results show more than two fold improvement when using depth images. In this analysis, we had the images taken under controlled lighting. In a practical setting, the intensity-based images would further suffer from lighting variations. In Figure 4, we provide images with two extreme lighting conditions. In the first, there is no light in the scene. In the second, we used 2 600-Watt lighting systems. Observe that the brightness images change tremendously from all black to bright, whereas the depth image does not change.

In Figure 5, we provide failure cases for the depth-based system. Most of these cases were either not part of the training set, or they belong to the extreme configurations of the occupant.

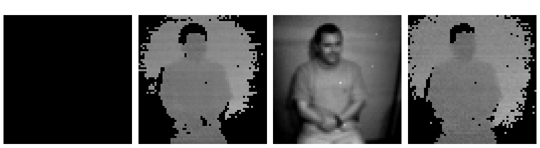

Figure 4. Images captured with no light and with extreme light conditions. The brightness image shows an extreme change, while the range image is not affected.

Figure 5. Some failure cases for the algorithm. These cases usually included extreme motion blur, saturation, or extreme configurations of the occupant. There were many similar cases, where the system made a correct classification.

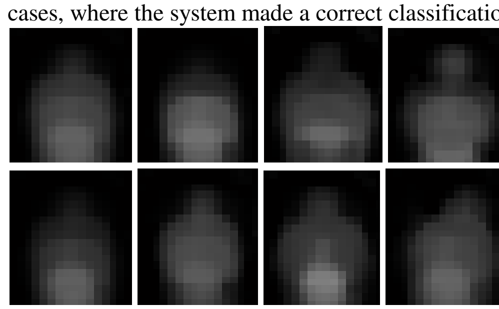

Figure 6. The first 8 Eigen Shapes. Observe that the shapes cover various adult, children, child seat and object cases.

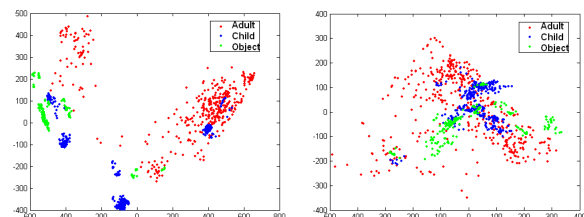

Figure 7. Distribution of the Eigen Shape projections on the first 4 components.

4.2. Eigen Shapes versus Knowledge-Based Features

As mentioned in Section 3.2, Eigen shapes and various other knowledge-based features were obtained to be used as features to the classification algorithm. In this analysis, we used depth images as input, and SVM with a kernel of exponential radial basis functions. First, we provide the first 8 principal components in Figure 6. Observe that the principal components capture various shapes, some of which belongs to a particular occupant type. Table 2 lists the results when different feature sets are used. The referenced feature sets are described in Section 3.2. The dimensions of the feature sets were 20, 12, 16, 12 and 36 for Eigen shapes, set 1, set 2, set 3, and set 4 respectively.

Our results show that the Eigen shapes, when used as features, provides better results compared to the knowledge-based features. This indicates that the principal directions of projections capture more relevant information compared to heuristics generated features. The Eigen shape provides an error rate of 1.63% while the best combination of features provides an error rate of 2.23%. It is also noted that the combination of the Eigen shapes and knowledge-based features provides the best results, with an error rate of 1.45%. This illustrates that the characteristics captured in the Eigen Shapes and the knowledge-based features somewhat complement each other. We contribute this to the following: The Eigen shapes capture the most commonly occurring shapes in the training set. The knowledge-based features, however, might help with the less frequent cases as well since they are created by human heuristics to resolve these cases.

Using Eigen shapes reduces the dimension of the image space optimally. This has two main advantages: First, the performance of the classification algorithm is improved as mentioned in the next section. Second, the classification algorithm runs faster with low-dimensional features. As such, a real-time system becomes possible.

4.3. SVM versus Other Classifiers

Finally, we provide results of our algorithm with various classifiers. Here, we used depth images as input, and Eigen shapes as our features. We applied SVM [19] with different kernels such as linear kernel, radial basis function(RBF(m)), exponential radial basis function (ERBF(m)), where m denotes the degree of the kernel function. For comparison purposes, we also applied linear discriminant analysis (LDA) [20], nearest neighbor classification (1-NN), and k-nearest classification with k=15 (15-NN), and k=8 (8-NN). Table 3 lists the results of this analysis.

First, it is observed that the linear classifiers such as SVM with linear kernel, or LDA did not work as well compared to the non-linear classifiers. To analyze this, we projected the Eigen Shape feature vectors of the training set onto first few basis directions (Figure 7). It is observed that the distribution of the training set is clearly non-linear. This is expected, since Eigen analysis is a linear transformation, and does not exploit the non-linearity's of complex cases as in our experiments. It is observed that the exponential radial basis functions provides the best results among all the classifiers.

The SVM are able to minimize the structural risk, given as the probability of misclassifying previously unseen data. More generally, using a learning method is necessary for a complex problem such as occupant classification where it is not quite possible to find the distinguishing components of the two classes due to the vast amount of configurations. Using SVM rather than hand-crafting classification heuristics, exploits all of the information in the training set optimally, and eliminates the guess work from the task of defining appropriate discrimination criteria. The performance of SVM, or other training based classifiers, depends heavily on the dimension of the system [12]. Using Eigen shapes, the dimension of the feature space is reduced in an optimal way, and the performance of the clssification system is improved.

It is common that the research in computer vision either focuses on the features or classifiers. It is usual that the feature vectors are fixed, and the research focuses on the classifiers, or vice versa. For a practical application, it is necessary that the performance optimization search space should be a combination of the feature extraction and the classifiers. The right combination of the two is the key to a successful system. In our application, the combination of a linear transformation for feature extraction with a non-linear classifier provides the best results.

6. Conclusions

In this paper, we illustrate the use of a time-of-flight depth sensor in an occupant classification system. We provide a solution that works with large amount of configurations of different occupant types, and under extreme lighting conditions, such as at night or under direct sun-light.

Our algorithm provides a solution based on Eigen shapes and Support Vector Machines. An input depth image is first segmented from the seat, and normalized for transformations of the camera. After normalization, every final image appears as if captured from the same virtual camera. Each normalized image is then processed to obtain Eigen shape features, as well as various heuristics based features. The features of the training set are used to learn a Support Vector Machine Classifier. The learnt classifier is applied on features of test cases to classify them into the appropriate occupant type.

The paper makes various contributions. First, it uses a novel depth sensing technique, and compares its results to traditional intensity based approach. Next, we provide a real-time and practical solution to the occupant classification problem. We applied a recognition framework, where we optimized the features and the classification algorithm and showed the superiority of our combination to several other choices. Our recognition framework also works with intensity images, although it wouldn't perform well especially under varying ambient conditions. Due to the combination of the depth sensor and the recognition framework, the system performs well under many configurations, and different lighting conditions. The current work has been applied on single images only. Future research will focus on combining the results from multiple frames in a layered network framework or using HMMs. We expect that the error rate can be improved from one in hundreds, to one in millions with such a framework.

References

[1] S.B. Gokturk, C. Tomasi, "3D Head Tracking Based on Recognition and Interpolation Using a Time-Of-Flight Depth Sensor", in *Proceedings of Computer Vision and Pattern Processing*, CVPR , pages 211-217, Washington DC, USA, 2004.

[2] "http://www.intellimec.com/products/ows.htm"

[3] "http://www.steadlands.com/oc.htm"

[4] M. E. Farmer, A. K. Jain, "Occupant Classification System for Automotive Airbag Suppression", *Conference on Computer Vision and Pattern Recognition*, IEEE, Madison, 2003, pp. 756-761.

[5] H. Kong, Q. Sun, W. Bauson, S. Kiselewich, P. Ainslie, R. Hammoud, "Disparity Based Image Segmentation For Occupant Classification", *Computer Vision and Pattern Recognition Workshop*, IEEE, Washington D.C., 2004.

[6] M. Turk and A. Pentland, "Eigen-faces for recognition", *Journal of Cognitive Neurosciences*, vol.3, no.1 77-86, 1991.

[7] V. Deepu, A. G. Ramakrishnan, "Principal Component Analysis for Online Handwritten Character Recognition", International Conference on Pattern Recognition (ICPR'04), IEEE, Cambridge, UK, pp. 327-330.

[8] M.Pontil, and A. Verri, "Support Vector Machines for 3D Object Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, Issue 6, pp. 637-646.

[9] S.B. Gokturk, C. Tomasi, B. Acar, C.F. Beaulieu, D. Paik, B. Jeffrey, J. Yee, S. Napel, A Statistical 3D Pattern Processing Method For Computer Aided Detection of Polyps in CT Colonography, in IEEE Transactions on Medical Imaging, Vol. 20, No 12, pages 1250-1260, December 2001.

[10] K. Jonsson, J. Kittler, Y. P. Li, and J. Matas, "Support vector machines for face authentication", *British Machine Vision Conference*, BMVC'99, pages 543-553, 1999.

[11] M.H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods", *Proceedings of the Fifth International Conference on Automatic Face and Gesture Recognition (FG 2002)*, Washington D. C., pp. 215-220.

[12] R. Bellman, "Adaptive Control Processes: A Guided Tour", *University Press*, Princeton, 1961.

[13] CMOS-compatible Three-dimensional Image Sensor IC, United States Patent 6,323,942, November 27, 2001.

[14] Methods for CMOS-compatible three-dimensional image sensing using quantum efficiency modulation: United States Patent 6,515,740, February 4, 2003.

[15] S.B. Gokturk SB, H. Yalcin, C. Bamji, ""A Time-of-flight Depth Sensor, System Description, Issues and Solutions", on IEEE Workshop on "Real-Time 3D Sensors and Their Use" in conjunction with IEEE Conference on Computer Vision and Pattern Recognition, CVPR, Washington DC, USA, 2004.

[16] I.T. Jolliffe, *Principal Component Analysis*, New York: Springer, 1986.

[17] A.G.Mamistvalov, n-dimensional moment invariants and conceptual mathematical theory of recognition n-dimensional solids, International Conference on Pattern Recognition (ICPR'90)., pp. 288 -290, 1990.

[18] V. Vapnik, *Statistical Learning Theory*, New York, 1998.

[19] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp.121-167, 1998.

[20] G.J.McLachlan, *Discriminant Analysis and Statistical Pattern Recognition,* John Wiley&Sons, Inc., 1992.