Embedding Vehicle Driver Face Poses on Manifolds

Hari C.V.*†‡, Praveen Sankaran*‡
*Department of Electronics & Communication Engineering
National Institute of Technology, Calicut, Kerala, India
†Department of Applied Electronics & Instrumentation Engineering
Rajagiri School of Engineering & Technology,
Kochi, Kerala, India
‡Email: haricv@gmail.com, psankaran@nitc.ac.in

***Abstract*—The number of deaths due to the motor vehicle accidents has increased drastically during the last few years. It is estimated that nearly 40 - 50 % of the accidents are due to the distractions of the driver. One of the methods to assess driver distraction is using head pose estimation. The unavailability of a standard driving database hampers the study and testing of driver distraction algorithms. In this paper we present the development and bench-marking of the Distracted Car Driver (DCD) database. The database contains real time driving videos of 12 different individuals, in varying environmental conditions. To bench-mark the database we use standard linear and nonlinear manifold techniques for data embedding.**

***Keywords*—Driver Distraction, Pose Estimation, Data Embedding, Discriminant Analysis.**

I. Introduction

Driver distraction is a leading cause of the accidents. One way to address this problem is by providing an alert to the vehicle occupants. The level of attention or distraction can be measured using the position and orientation of the driver head. There are many driver distraction observational studies and head pose estimation methods carried out all over the world during the past few years. Sayer *et.al* [1] observed that 34 percent of the drivers involved in a secondary task, which will cause an accident. The Highway Safety Research Center at the University of North Carolina found that drivers spent 15.3 percent of the time for conversation with passengers and 14.5 percent engaged in some other activities of the total driving time [2]. The driver drowsiness or sleepiness is also one of the major reason for road accidents especially in the late nights [3], [4]. Sigari et.al [5] give a detailed review of driver face monitoring system for fatigue and distraction detection. The face monitoring systems for distraction detection extract the different features of face and eyes like head orientation, yawning, head nodding, gaze direction, eye distances, eye blink rate, blink speed . . . etc. Head rotation based on face template matching and eye region monitoring are used for fatigue and driver distraction monitoring in [6], [7]. The main disadvantage of this method is the face tracking method which is inaccurate and very computationally complex.

In the literature, it is seen that face pose estimation has an important role in driver distraction monitoring. Principal Component Analysis (PCA) [8] is used for face pose estimation by projecting the data samples into its principal components [9],[10]. Linear Discriminant Analysis (LDA) [11] is another linear dimensionality reduction method and it is also used to obtain the discriminant features for pose estimation in[12],[13]. If the number of data points for training is small, PCA performance is better than LDA and it is less sensitive to different training sets [14]. Isometric Feature Mapping (ISOMAP) [15], Locally Linear Embedding (LLE) [16] and Laplacian Eigenmaps (LE) [17] are the most common non linear dimensionality reduction methods used for pose estimation [18], [19].

All the above methods test using standard face pose databases [20],[21]. These databases are developed in controlled environments and provide manually annotated face pose angle information. In a real world scenario as in vehicles, these databases fall short. There is a limited scope to control the environment and pose angles, which results in manually annotating the angles a daunting task. Hence, we develop a driving database (Distracted Car Driver (DCD)), which would aid in development and testing of algorithms to estimate constraint free face poses and in turn develop driver distraction analysis algorithms.

II. Database Creation

Video of 12 individuals (9 male and 3 female) was captured by playing a camera on the dashboard. To ensure the variations in the light, the video capturing was done in three different timings - morning, noon and at evening. Separate scenarios in vehicle environments was considered where one scenario had co-passengers and another where there were none. This was done to simulate the distractions due to co-occupants of the vehicle. The video was finally separated into frames.

Five different classes were created and named as follows. Non - Distracted (approximately the pose angles are varying from -15° to $+15^{\circ}$), Small Distraction to the left side (approximately the pose angles are varying from -30° to -16°), High Distraction to the left side (approximately the pose angles are varying from -90° to -31°), Small Distraction to the right side (approximately the pose angles are varying from $+16^{\circ}$ to $+30^{\circ}$) and High Distraction to the right side (approximately the pose angles are varying from $+31^{\circ}$ to $+90^{\circ}$)

Five different persons manually separated these images into any one these five distinct classes and these five sets are kept separately. If a non distracted frame is present in the non distracted class at least thrice out of these five sets, it will be included in the non distracted class. Similarly, we evaluated all frames and placed into the corresponding classes. This final

978-1-5386-3864-4/17/$31.00

database is named as Distracted Car Driver (DCD) database . The sample images from this DCD database are as shown in Fig.1.

III. Dimensionality Reduction Methods

It is very difficult to understand the data if the dimensions are more than two or three, and this is the main reason to obtain the structure of the data by low dimensional representation. If the dimension of the data is reduced, the data can be easily processed than in higher dimensions [22]. A brief description about different dimensionality reduction techniques follows.

A. Linear Dimensionality Reduction Methods

Linear techniques are simple geometric interpretations and has attractive computational properties [23]. Principal Component Analysis (PCA) [8] and Linear Discriminant Analysis (LDA) [11], [24]. are two popular linear techniques for dimensionality reduction.

PCA gives the most accurate data representation in a lower dimensional space. It identifies the variation in the data based on identifying the Principal Components in the data. Principal Components are in the directions of maximum variability in the data. PCA transforms the data into a 1-D subspace which minimizes the projection error. The best projection line lies in the direction of largest variance and this is called the first principal component. The second principal component in the direction of next highest variability and it is orthogonal to the first principal component. The subsequent principal component in decreasing order in the direction of highest variability. LDA tries to maximize the between class separability and to minimize the within-class variability, and so ensuring the maximal separability.

B. Non Linear Dimensionality Reduction Techniques

Closely varying data points creates a smooth varying manifold. Linear techniques can not capture the intrinsic nonlinearities. The nonlinear techniques try to preserve the the neighborhood information. Commonly used nonlinear techniques are Isometric Feature Mapping, Locally Linear Embedding, and Laplacian Eigenmaps.

Tenenbaum *et.al* [15] proposed the isometric feature mapping (ISOMAP) for nonlinear dimensionality reduction. After identifying the K nearest neighbours of each data point based on the euclidean distance between them, this neighbourhood points are represented on a weighted graph. Dijkstra's algorithm [25] is used to obtain the geodesic distances between the points and classical multidimensional scaling is applied onto the geodesic distance matrix. The eigenvectors obtained are used for data embedding. Roweis *et.al* [26] proposed the locally linear embedding (LLE). LLE kernel is formed using barycentric coordinates and eigen decomposition of these LLE kernels is performed. The obtained eigenvectors are used for the data embedding. Belkin *et.al* [17] proposed the laplacian eigenmaps method for data embedding. The neighborhood graph is approximated to Laplace Beltrami operator [27]. The weight matrix is formed using the Leigs algorithm [28] and the kernel is constructed. The eigen vector obtained after eigen decomposition of these kernel is used for data embedding.

IV. Linear Techniques Based Pose Estimation for Driver Distraction Monitoring

Initially the entire data set is divided into different classes for training. Clustering methods can be used for automatic class formation for training [12],[13]. The automated class formation for clustering reduces the effort of manual process. The classes which form after clustering should be based on the pose angles and hence will be pose dependent. A conceptual drawing of this pose dependent class formation using clustering is as shown in Fig.4. The data embedding of PCA and LDA in top two dimensions are as shown in Fig. 2 and 3 respectively.

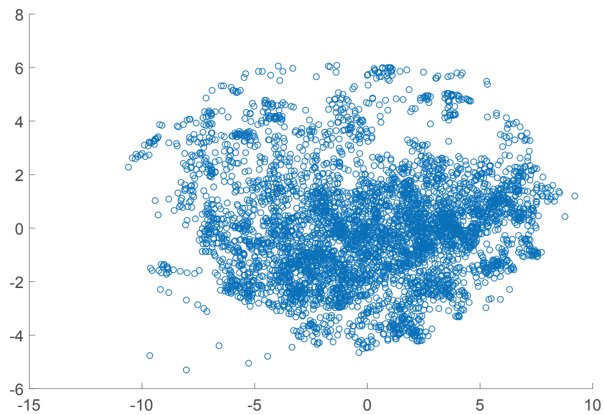


Fig. 2: PCA - Data Embedding.

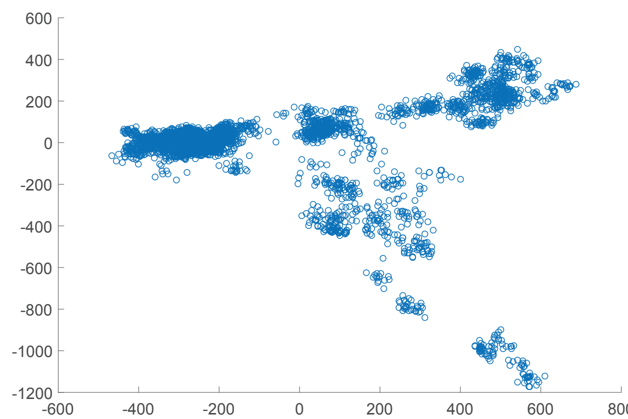
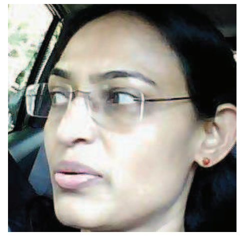
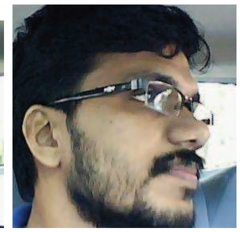


Fig. 3: LDA - Data Embedding.

Initially the entire dataset is clustered into various classes using K means [29],[30] clustering algorithm. Then, PCA and LDA applied on these classes and nearest neighbour algorithm is used for classification. In the case of linear dimensionality reduction methods, a direct projection matrix is available for data embedding. In both PCA and LDA, all training data points and test data point(s) are projected using this projection matrix. The euclidean distance between the test data point(s) and training data points is calculated in the projected space. The training data point with the shortest distance to the test



(a) High Distraction to the Left Side (b) Small Distraction to the Left Side



(c) Non Distracted (d) Small Distraction to the Right Side (e) High Distraction to the Right Side

Fig. 1: Distracted Car Driver (DCD) Database - Sample Images

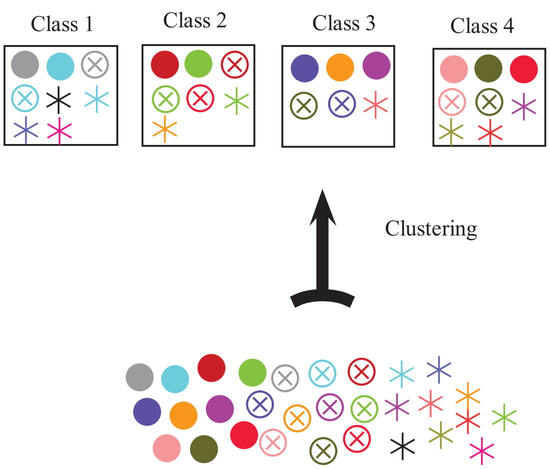


Fig. 4: Pose Dependent Class Formation Using Clustering - Conceptual Drawing.

data point is considered as the winner. The class of the winning image is taken as the detected class.

A. Experimental Analysis

Leave one out testing strategy was used for performance evaluation. In this testing procedure, the data points used for training are not used for testing. Images of 11 individuals are used for training and the images of person left out are used for testing. This is iterated till the images of all individuals are used for testing. If the test image is correctly classified into its own class, it is taken as true and if it is not, it is taken as false. The average accuracy is calculated by dividing the total number of true values with the total number of testing.

The data points used for testing is not used for clustering (leave-one-out testing). We varied the number of clusters and iterated till all data points used for testing. We tested and evaluated this method using K means clustering and the result of this analysis is as given in Table I. From Table I, it is very clear that the performance of LDA outperforms PCA in pose estimation.

TABLE I: Driver Distraction Monitoring - Accuracy in Pose Estimation

Number of Classes		**10**	**25**	**50**	**100**	**150**	**181**
Accuracy	**PCA**	45.36	50.23	58.36	65.12	68.25	78.86
	LDA	71.69	78.89	84.21	86.76	87.07	88.80

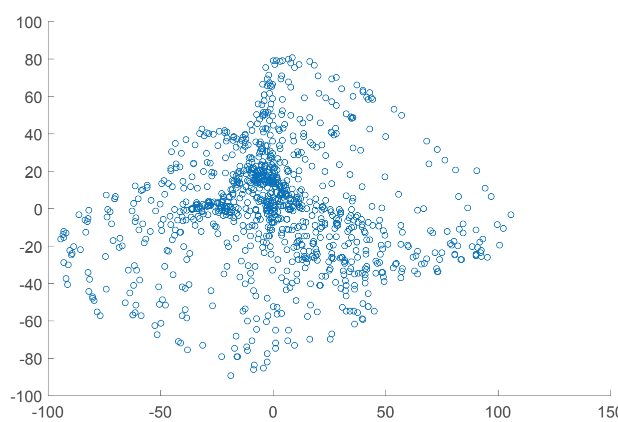


Fig. 5: ISOMAP - Data Embedding.

V. Non Linear Techniques for Data Embedding

Non linear techniques - ISOMAP, LLE and LE was used for data embedding. This techniques do not provide a projection matrix for other test data points. These techniques do not

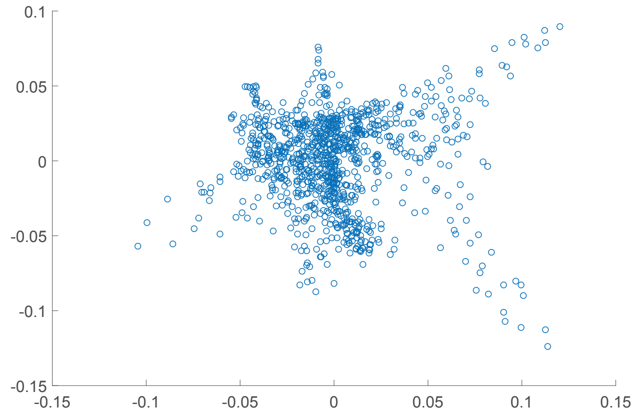


Fig. 6: LLE - Data Embedding.

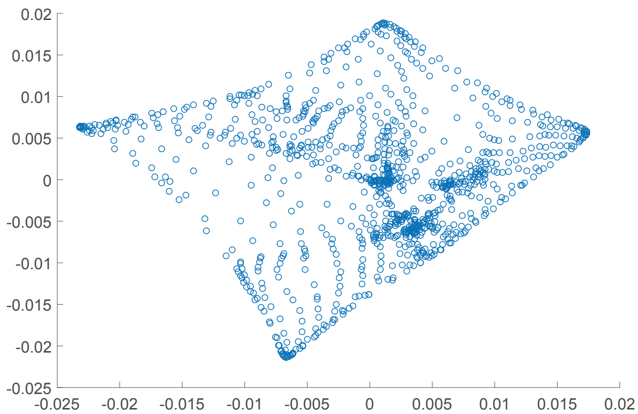


Fig. 7: LE - Data Embedding.

provide a projection matrix and so the out of sample extension is not possible. So these techniques are used for observing the non-linearity in the database. The data embedding of these methods in top two dimensions are as shown in Figure 5, 6 and 7.

VI. Conclusion

The unavailability of a standard driving database was a major problem for driver distraction monitoring. We created a database (DCD) for driver distraction monitoring with multiple people in different environmental conditions. Varied results, with low accuracy points to the toughness and usability of the database for testing algorithms tackling real world pose estimation problems. During the data collection and embedding steps we noticed that the database could be made more robust by syncing video of vehicle occupant with vehicle frontal video. The database could be made more comprehensive by logging more hours of driving video.

References

[1] J. R. Sayer, J. M. Devonshire, and C. A. Flannagan, "The effects of secondary tasks on naturalistic driving performance," University of Michigan, Transportation Research Institute, Tech. Rep., 2005.

[2] J. Stutts, J. Feaganes, D. Reinfurt, E. Rodgman, C. Hamlett, K. Gish, and L. Staplin, "Driver's exposure to distractions in their natural driving environment," *Accident Analysis & Prevention*, vol. 37, no. 6, pp. 1093–1101, 2005.

[3] D. FATIGUR, "Road accidents. a literature review and position paper"," 2001.

[4] S. McKernon, *A literature review on driver fatigue among drivers in the general public*. Land Transport New Zealand, 2009.

[5] M.-H. Sigari, M.-R. Pourshahabi, M. Soryani, and M. Fathy, "A review on driver face monitoring systems for fatigue and distraction detection," 2014.

[6] M.-H. Sigari, M. Fathy, and M. Soryani, "A driver face monitoring system for fatigue and distraction detection," *International journal of vehicular technology*, vol. 2013, 2013.

[7] M. H. Sigari, "Driver hypo-vigilance detection based on eyelid behavior," in *Advances in Pattern Recognition, 2009. ICAPR'09. Seventh International Conference on*. IEEE, 2009, pp. 426–429.

[8] K. Person, "On lines and planes of closest fit to system of points in space. philiosophical magazine, 2, 559-572," 1901.

[9] S. J. McKenna and S. Gong, "Real-time face pose estimation," *Real-Time Imaging*, vol. 4, no. 5, pp. 333–347, 1998.

[10] S. Srinivasan and K. L. Boyer, "Head pose estimation using view based eigenspaces," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, vol. 4. IEEE, 2002, pp. 302–305.

[11] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[12] C. Hari and P. Sankaran, "A clustered locally linear approach on face manifolds for pose estimation," *Pattern Analysis and Applications*, pp. 1–10, 2016.

[13] ——, "Face pose estimation for driver distraction monitoring by automatic clustered linear discriminant analysis," in *Vehicular Electronics and Safety (ICVES), 2014 IEEE International Conference on*. IEEE, 2014, pp. 100–105.

[14] A. M. Martínez and A. C. Kak, "Pca versus lda," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, no. 2, pp. 228–233, 2001.

[15] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[16] L. K. Saul and S. T. Roweis, "An introduction to locally linear embedding," *unpublished. Available at: http://www. cs. toronto. edu/~ roweis/lle/publications. html*, 2000.

[17] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[18] B. Raytchev, I. Yoda, and K. Sakaue, "Head pose estimation by nonlinear manifold learning," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 4. IEEE, 2004, pp. 462–466.

[19] V. N. Balasubramanian, S. Krishna, and S. Panchanathan, "Person-independent head pose estimation using biased manifold embedding," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, pp. 1–15, 2007.

[20] D. Little, S. Krishna, A. John Jr, and S. Panchanathan, "A methodology for evaluating robustness of face recognition algorithms with respect to variations in pose angle and illumination angle." in *ICASSP (2)*. Citeseer, 2005, pp. 89–92.

[21] N. Gourier, D. Hall, and J. L. Crowley, "Estimating face orientation from robust detection of salient facial features," in *ICPR International Workshop on Visual Observation of Deictic Gestures*. Citeseer, 2004.

[22] L. Cayton, "Algorithms for manifold learning," *Univ. of California at San Diego Tech. Rep*, pp. 1–17, 2005.

[23] J. P. Cunningham and Z. Ghahramani, "Linear dimensionality reduction: Survey, insights, and generalizations," *Journal of Machine Learning Research*, vol. 16, pp. 2859–2900, 2015.

[24] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *Pattern*

Analysis and Machine Intelligence, IEEE Transactions on, vol. 19, no. 7, pp. 711–720, 1997.

[25] S. Skiena, "Dijkstras algorithm," *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica, Reading, MA: Addison-Wesley*, pp. 225–227, 1990.

[26] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[27] L. Sario, M. Nakai, C. Wang, and L. O. Chung, "Laplace-beltrami operator," *Classification Theory of Riemannian Manifolds: Harmonic, quasiharmonic and biharmonic functions*, pp. 12–25, 1977.

[28] J. Wang, *Geometric structure of high-dimensional data and dimensionality reduction*. Springer, 2011.

[29] E. Gose, "Pattern recognition and image analysis," 1997.

[30] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14. Oakland, CA, USA., 1967, pp. 281–297.