NULL-SPACE REPRESENTATION FOR VIEW-INVARIANT MOTION TRAJECTORY CLASSIFICATION-RECOGNITION AND INDEXING-RETRIEVAL

Eser Ustunel, Xu Chen, Dan Schonfeld and Ashfaq Khokhar

{eustun2, xchen27, dans, ashfaq}@uic.edu

ABSTRACT

This paper presents a novel classification/ retrieval system for motion events based on a perfect view invariant representation of motion trajectories and a linear classifier algorithm. Specifically, Null Space Invariant (NSI) matrix representation for motion trajectories has been derived. The proposed view invariant representation based on the NSI operator is invariant to affine transformations and preserves the null space matrix. We use principal component null space analysis (PCNSA) for indexing and classification of the NSI operator for recognition and retrieval of motion events. We rely on PCNSA to determine the distance of the query trajectory to the centroid of each class, which is a statistical information vector in the PCNSA algorithm representing the corresponding motion class. Our results show that NSI provides a powerful approach to motion event recognition and retrieval that is invariant to affine transformations due to camera motion.

***Index Terms*—** Trajectory Retrieval, Affine View Invariants, Principal Component Analysis, Object Recognition and Classification.

1. INTRODUCTION

Within the last several years, object motion trajectory-based recognition has gained significant interest in diverse application areas including sign language gesture recognition, Global Positioning System (GPS), Car Navigation System (CNS), animal mobility experiments, sports video trajectory analysis and automatic video surveillance [1]. Psychological studies show that human beings can easily discriminate and recognize an object's motion pattern even from large viewing distances or poor visibility conditions where other features of the object vanish.

The development of accurate activity classification and recognition algorithms in multiple view situations is still an extremely challenging task. Object trajectories captured from different view-points lead to completely different representations, which can be modeled by affine transformation approximately. To get a view independent representation, the trajectory data is represented in an affine invariant feature space.

A new invariant algorithm for Structure From Motion (SFM) problem is proposed in [2]. It uses invariant property of group action on a vector space to eliminate the camera pose parameters in the calculations. That enables robust solutions for SFM. Although the invariant algorithm is promising to use in recognition and classification problems, the requirement for solving the invariant equations by Levenberg-Marquardt algorithm introduces high computational complexity. In [1] two different affine invariant representations for motion trajectories namely: Curvature Scale Space (CSS) and Centroid Distance Functions (CDF) have been used for trajectory recognition and classification. However, the performance of these techniques has been limited in the presence large number of objects or classes.

In this paper, we introduce a simple but highly efficient view invariant representation based on Null Space Invariant (NSI) matrix [3]. As far as we know, this is the first use of Null space in motion-based classification/retrieval applications.

Indexing and classification of the NSI operator is obtained by extracting features of the null space representation using Principal Components Null Space Analysis (PCNSA), which provides an efficient analysis tool when different classes may have different non-white noise covariance matrices [4]. Dimensionality reduction for indexing of the NSI is achieved by first performing Principal Components Analysis (PCA) as part of PCNSA. Classification is performed in PCNSA by determining the ith class M_i-dimensional subspace by choosing the M_i eigenvectors that give the smallest intra-class variance. The M_i-dimensional space is referred to as the Approximate Null Space (ANS). A query is classified into the class if its distance to the class mean in ANS space is lowest among all the other classes.

The rest of the paper is organized as follows: The new invariant representation is presented in Section 2. The PCNSA algorithm is discussed in Section 3. Section 4 provides details of the data set and simulation results. Finally, in Section 5, we present a brief summary of our results.

2. NULL SPACE INVARIANT

A fundamental set of 2-D affine invariants for an ordered set of n points in R^2 (not all collinear) is expressed as an n-3 dimensional subspace, H^{n-3}, of R^{n-1}, which yields a point in the 2n-6 dimensional Grassmannian Gr_R(n-3,n-1), which also shows number of invariants is 2n-6 in 2-D.

Null Space Invariant (NSI) of a trajectories matrix (each row in the matrix corresponds to the positions of a single object over time) is introduced as a new and powerful affine invariant space to be used for trajectory representation. This invariant, which is a linear subspace of a particular vector space, is the most natural invariant and is definitely more general and more robust than the familiar numerical invariants. It does not need any assumptions and after invariant calculations it conserves all the information of original raw data.

Let $Q_i = (x_i, y_i)$ be a single 2-D point, $i = 0, 1, \ldots, n-1$, among n ordered non-linear points in R^2, representing a trajectory. Consider the following arrangement of the n 2-D points in a $3 \times n$ matrix M:

$$M = \begin{pmatrix} x_0 & x_1 & \ldots & x_{n-1} \\ y_0 & y_1 & \ldots & y_{n-1} \\ 1 & 1 & \ldots & 1 \end{pmatrix} \tag{1}$$

(n-3)-dimensional linear subspace H^{n-3} can be associated to a 2-D trajectory whose features set is $Q_0, Q_1, \ldots, Q_{n-1}$:

$$H^{n-3} = \{q = (q_0, q_1, \ldots, q_{n-1})^T, i.e. Mq = (0,0,0)^T\} \tag{2}$$

Since at least one determinant of 3×3 minor of M is not zero because of non-linear feature points, H^{n-3} has a dimension of n-3. The attractive property of the linear subspace is that it does not change when it undergoes any of the affine transformations. We use this new invariant of the trajectory matrix M to represent each trajectory. Moreover,

$$H^{n-3} \subset R^{n-1} = \{q = (q_0, q_1, \ldots, q_{n-1})^T \in R^{n-1}, \\ and \sum_{i=0}^{n-1} q_i = 0\} \tag{3}$$

which produces (n-3)-planes in (n-1)-space, $Gr_R(n-3, n-1)$. $Gr_R(n-3, n-1)$ is a well understood manifold of dimension 2n-6, which is the number of invariants associated to the matrix M. H^{n-3} is spanned by the vectors $v_i = (q_0^i, q_1^i, \ldots, q_{n-1}^i)^T$, $i = 3, 4, \ldots, n-1$, where

$$q_0^i = -det \begin{pmatrix} x_1 & x_2 & x_i \\ y_1 & y_2 & y_i \\ 1 & 1 & 1 \end{pmatrix} / det \begin{pmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ 1 & 1 & 1 \end{pmatrix} \tag{4}$$

$$q_1^i = det \begin{pmatrix} x_0 & x_2 & x_i \\ y_0 & y_2 & y_i \\ 1 & 1 & 1 \end{pmatrix} / det \begin{pmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ 1 & 1 & 1 \end{pmatrix} \tag{5}$$

$$q_2^i = -det \begin{pmatrix} x_0 & x_1 & x_i \\ y_0 & y_1 & y_i \\ 1 & 1 & 1 \end{pmatrix} / det \begin{pmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ 1 & 1 & 1 \end{pmatrix} \tag{6}$$

$$q_i^i = -1 and q_i^j = 0 for j = 3, 4, \ldots, i-1, i+1, \ldots, n-1 \tag{7}$$

This is an elegant and simple method to find the null space of a trajectory matrix M, which is also the most promising invariant of a trajectory. It is important to note that not every basis representation of the null space can be used to form an invariant matrix for a trajectory. For example, using the singular value decomposition (SVD) to form an orthogonal basis of the null space would not guarantee an affine invariant representation. Now each trajectory is represented now with a matrix $NSI_{n \times (n-3)}$, which is composed of v_i columns, from the trajectory matrix $M_{3 \times n}$. While is an affine invariant matrix, this representation increases the dimensionality significantly, from 2n to $n \times (n-3)$. We reduce the dimensionality to a comparably small but effective value such as n by using PCA.

Elements along the trajectory obtained from different camera positions, e.g. moving camera, can be thought of as point-wise affine transformations that are applied on the feature points of a trajectory. It can be shown that affine transformations of the trajectory at each point results in an overall affine transformation of the entire trajectory. Let A_i denote a matrix in SO(2) , i.e. a matrix with $det(A_i) = 1$ and $A_i^T A_i = I_2$, where I_2 is a 2×2 identity matrix. The matrix A_i rotates any feature point Q_i to a new point in R^2 as $Q_i' = A_i Q_i$, for $i = 0, 1, \ldots, n-1$. We can now verify that $\bar{A} = \begin{pmatrix} A_0 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & A_{n-1} \end{pmatrix}$ is a $2n \times 2n$ matrix in SO(2n),i.e.a matrix such that $det(\bar{A}) = 1$ and $\bar{A}^T \bar{A} = I_{2n \times 2n}$.

3. PRINCIPAL COMPONENT NULL SPACE ANALYSIS

Once we have generated the affine invariant representation provided by the null space operator, $NSI_{n \times (n-3)}$, we can rely on numerous methods for indexing and classification. We choose a method for dimensionality-reduction and classification based on PCNSA. Notice that the term null space used in PCNSA is meant that the Approximate Null Space (ANS) used for representation of each class is formed from the minimal eigenvectors within the class and thus minimizes the intra-class variance. However, this process is not intended to capture the null operator and is unrelated to the null space invariant proposed in Section 2.

First, NSI is converted to $P = n(n-3)$ column vector Y_p which is assumed in class C_i and has Gaussian distribution as $Y | \{Y \in C_i\} \sim N(\mu_{full,i}, \Sigma_{full,i})$, where $\mu_{full,i}$ is the class conditional mean and $\Sigma_{full,i}$ is the class conditional covariance matrix. To decrease the high dimensionality due to $NSI_{n \times (n-3)}$, we perform Principal Component Analysis (PCA), which removes the noise-only directions and retain the directions that yield large inter-class variance. PCA takes the L leading eigenvectors of covariance matrix, Σ_{full}, of the entire data taken from all classes. The total scatter matrix, Σ_{full}, can be written $\Sigma_{full} = \Sigma_{full,w} + \Sigma_{full,b}$ where $\Sigma_{full,w}$ is within class covariance matrix and $\Sigma_{full,b}$ between class covariance matrix i.e.

$$\Sigma_{full,w} = \frac{1}{C}\sum_{i=1}^{C}\frac{1}{K}\sum_{k=1}^{K}(Y(i,k)-\mu_{full,i}) (Y(i,k)-\mu_{full,i})^T , \quad (8)$$

$$\Sigma_{full,b} = \frac{1}{C}\sum_{i=1}^{C}(\mu_{full,i}-\mu_{full})(\mu_{full,i}-\mu_{full})^T , \quad (9)$$

where i is for class index and k is for trajectory index in the class. It is assumed that there are C classes in the system and each class has K trajectories.
PCA gives the L-dimensional projection matrix $(W_{PCA})_{P\times L}$ and the projections into the PCA space are

$$(X)_{L\times 1} = W_{PCA}^T(Y-\mu_{full}) \sim N(\mu_i,\Sigma_i) \quad (10)$$
$$(\mu_i)_{L\times 1} = W_{PCA}^T(\mu_{full,i}-\mu_{full}) \quad (11)$$
$$(\Sigma_i)_{L\times L} = W_{PCA}^T\Sigma_{full,i}W_{PCA} \quad (12)$$

After projections, in the PCA space PCNSA finds for each class i an M_i dimensional subspace along which the class's intra-class variance is smallest. This subspace is referred to as the Approximate Null Space (ANS) denoted as N_i since the lowest eigenvalues' corresponding eigenvectors are taken. That means we choose the lowest noise variance directions as for ANS.
Assumptions:

1. If $\lambda_{max,i}$ and $\lambda_{min,i}$ are the maximum and minimum eigenvalues of Σ_i, there should be a threshold number such that $\lambda_{max,i}/\lambda_{min,i} > \delta_1$. This guarantees Approximate Null Space (ANS) for ith class.

2. There should be an another threshold number such that $|(\mu_i-\mu_j)^Te_i| > \delta_2\|\mu_i-\mu_j\|$, where e_i is any column of N_i. This guarantees that any class i is linearly separable form other class j.

To get better solutions $\delta_1 \times \delta_2$ multiplication should be high i.e. $\delta_1 = 10^7, \delta_2 = 10^{-4}$ can work. **PCNSA Algorithm:**

1. **Obtain PCA Space:** Evaluate the total covariance matrix Σ_{full}, then apply PCA to the Σ_{full} to find $(W_{PCA})_{P\times L}$, whose columns are the L leading eigenvectors.

2. Project the data vectors, class means and class covariance matrices into the corresponding data vectors, class means, and class covariance matrices in the PCA space.

3. Obtain ANS: Find the approximate null space $(N_i)_{L\times M_i}$, for each class i by choosing M_i smallest eigenvalues' corresponding eigenvectors.

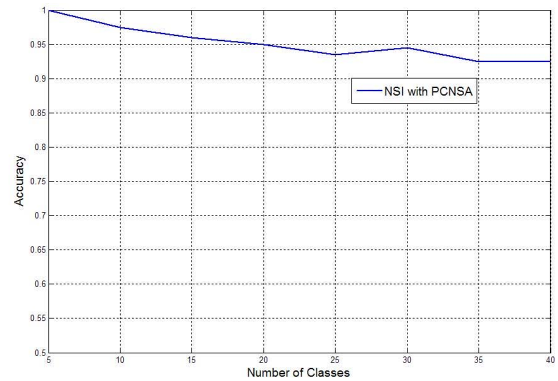

Fig. 1. Accuracy for motion trajectory classification with an increasing number of classes.

4. **Obtain Valid Classification Directions in ANS:** Say $N_i = (e_{i,1}|e_{i,2}|\ldots|e_{i,M_i})_{L\times M_i}$. If any direction, e_i satisfies $|(\mu_i-\mu_j)^Te_i| > \delta_2\|\mu_i-\mu_j\|$, this direction is said valid direction and used to build valid ANS, $W_{NSA,i}$.

5. **Classification:** PCNSA finds distances from a query trajectory to all classes

$$d_i(X) = \|W_{NSA,i}(X-\mu_i)\| . \quad (13)$$

We choose the smallest distance to a class for classification of X.

4. EXPERIMENTAL RESULTS

In order implement and evaluate the proposed classification and retrieval system, we have used trajectories from the the Australian Sign language (ASL) data set obtained from University of California at Irvine's Knowledge Discovery in Databases archive [5]. The trajectories in the data set are obtained by registration of the hand coordinates at each successive instant of time by using a Power Glove interfaced to the system. In our simulations, we used 40 different classes representing signing of 40 different words in the data set. Each class has 69 trajectories recorded at different instances.

Since in real life trajectories in a class may have different lengths, we normalize the length by taking the Fourier Transform and choosing the biggest n=32 coefficients and then taking the Inverse Fourier Transform so that all the trajectories are of size 32 before invariant matrix calculations. The invariant matrix, , for each trajectory is certainly robust to affine transformation such that it always preserves its values against rotation or translation operations. For all the simulations $\delta_1 = 10^7$, $\delta_2 = 10^{-4}$ as thresholds and $L = 32$ in PCNSA. Although we did not show here, we note that increasing L gives better results.

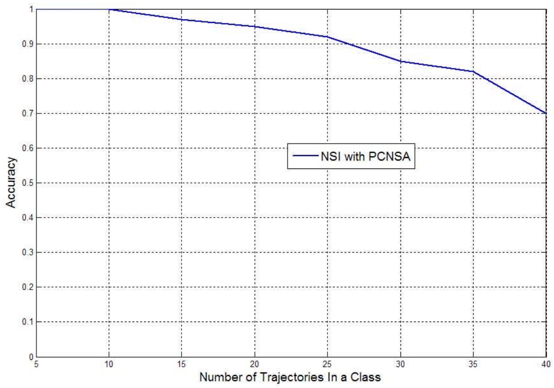

Fig. 2. Accuracy for motion trajectory classification with an increasing number of trajectories within each class.

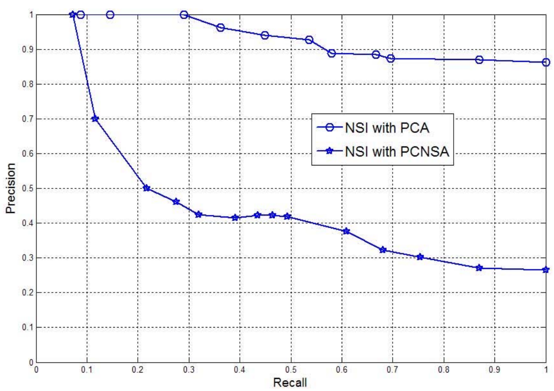

Fig. 3. Precision-Recall metric for motion trajectory retrieval using PCNSA on NSI.

Figure.1 depicts accuracy of the proposed classification system versus number of classes. There are $K = 20$ trajectories in each class word. Simulation results show that our system preserves its efficiency even for higher number of different classes.
Figure.2 depicts accuracy values versus increase in the number of trajectories within a class. There are $C = 20$ classes in the system. Simulation results show that our system performance deteriorates slightly for high number of trajectories in a class. This problem can be resolved by using a hierarchical representation, where we first separate all the trajectories in the system into a small number of classes, then repeatedly divide each class into smaller classes until each class has a sufficiently small number of trajectories. We can now use the null space representation for each class in the nested hierarchical tree structure to obtain a scalable representation whose performance is robust as the number of trajectories in the system increases. Figure.3 shows Precision vs. Recall curves for indexing and retrieval problem by using 40 classes, each class having 20 trajectories. For retrieval problems, we compute the distance of the query trajectory to any other trajectory using PCNSA on NSI as $D(X_i, Y) = \|W_{NSA,i}(X_i - Y)\|$, where Y is the query trajectory. This distance is then used to find α nearest trajectories, where α is a user specified parameter. There are two curves in Figure 3, one is with using PCA on NSI directly, where PCA is basically used for dimension reduction. As it can be seen from Figure 3 that the result of using PCNSA on NSI is much superior to the one using PCA on NSI directly.

5. CONCLUSION

We demonstrated the enormous potential of the NSI operator as a powerful view-invariant representation for recognition and retrieval. The computational complexity of NSI is very low and it possesses several important features, e.g. generality, robustness, and preserves information of the original data when computing invariants. We have also shown that the increased dimensionality of the NSI matrix representation can be effectively reduced by PCA. Moreover, any classification algorithms can be used based on PCA representation of the NSI operator. We further demonstrated the performance of classification of motion trajectories based on NSI and PCNSA in recognition and retrieval applications.

6. REFERENCES

[1] F.I. Bashir, A.A. Khokhar, and D. Schonfeld. ”Real-time affine-invariant motion trajectory-based retrieval and classification of video sequences from arbitrary camera view,” ACM Multimedia Systems Journal, Special Issue on Machine Learning Approaches to Multimedia Information Retrieval, vol. 12, pp. 45-54, 2006.

[2] P. L. Bazin and M. Boutin, ”Structure from Motion: a new look from the point of view of invariant theory,” SIAM Journal on Applied Mathematics, 64(4):1156-1174, 2004.

[3] P. F. Stiller, C. A. Asmuth, and C. S. Wan, ”A General Theory of Single View Recognition - the Affine Case - with Applications to Indexing Image Databases for Content Based Retrieval,” International Journal of Computer Vision 10/94.

[4] N. Vaswani, R. Chellappa, ”Principal Components Null Space Analysis for Image and Video Classification,” IEEE Trans. Image Processing, July 2006.

[5] S. Hettich, S. D. Bay: The UIC KDD Archive [http://kdd.ics.uci.edu]. University of California, Department of Information and Computer Science, Irvine, California, 1999.