Robust Null Space Representation and Sampling for View-Invariant Motion Trajectory Analysis

Xu Chen, Dan Schonfeld and Ashfaq Khokhar
Department of Electrical and Computer Engineering, University of Illinois at Chicago
xchen27, dans, ashfaq@uic.edu

Abstract

In this paper, we propose a novel robust retrieval and classification system for video and motion events based on null space representation. In order to analyze the robustness of the system, the perturbed null operators have been derived with the first order perturbation theory. Subsequently, the sensitivity of the null operators is discussed in terms of the error ratio and the SNR respectively. Meanwhile, the normwise bounds and componentwise bounds based on classical matrix perturbation theory are presented and discussed. Given the perturbation, uniform sampling are proposed for the convergence of the SNR and Poisson sampling are proposed for the convergence of the error ratio in the mean sense by choosing the rate parameter the same order as the number of samples. The simulation results are provided to demonstrate the effectiveness and robustness of our system in motion event indexing, retrieval and classification that is invariant to affine transformation due to camera motions.

1. Introduction

In recent years, object motion trajectory-based recognition has aroused significant interest, such as sign language data measurements, Car Navigation System(CNS), in sports video trajectory analysis and automatic video surveillance. An object trajectory captured from different view-points leads to completely different representations, which can be modelled as affine transformation approximately. To get a view independent representation, the trajectory data is represented in an affine invariant feature space. This paper addresses the important question of how to perform video retrieval and classification using motion trajectories when the query and video sequences in the database are taken from cameras with different view or possibly from moving cameras.

In [3], the mathematical form of the representation of the null space invariants has been derived. In this paper, we will rely on the theoretical formulation of the null space invariants to demonstrate its enormous potential in computer vision and related fields. Specifically, we will demonstrate the invariance of the null space representation of motion trajectories of moving objects to camera orientation and movement. We will subsequently investigate the robustness of the proposed approach to view invariance based on null space representation. We will evaluate the sensitivity of trajectory analysis to noisy data by using tools from perturbation theory. We will determine the performance of the null space representation in terms of both the error ratio and SNR. We will finally derive the optimal sampling strategy that minimizes the sensitivity to noise perturbation and thus ensures the robustness of null space representation. Moreover, we will demonstrate that the optimal sampling strategy required to maximize the SNR is provided by uniform sampling. We therefore observe that the null space representation is particularly suitable for the invariant representation of computer vision systems since they traditionally rely on uniform sampling for data acquisition (e.g. uniform sampling of the temporal dimension is used to capture video signals).

1.1. Null space representation

A fundamental set of 2-D affine invariants for an ordered set of n points in \( R^2 \) (not all colinear) 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 \( Gr_B(n-3, n-1) \), a manifold of dimension 2n-6. Null space invariant (NSI) of a trajectory 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 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 the original raw data. Let \( Q_i = (x_i, y_i) \) be a 2-D point for \( i=0,1,...,N-1 \) N ordered non-linear points in \( R^2 \). Consider
The invariants of the trajectory are defined as:

\[ H^{N-3} = \{ q = (q_0, q_1, \ldots, q_{N-1})^T \in \mathbb{R}^{N-1} \mid Mq = (0, 0, 0)^T \} \]  

2. Perturbation analysis

2.1. Perturbation analysis of null spaces

Perturbation analysis is an important mathematical method which is widely used for analyzing the sensitivity of linear systems by adding a small term to the mathematical description of the exactly solvable problem. Let’s assume the noise matrix \( Z \) as:

\[
\begin{pmatrix}
\epsilon_{x,0} & \epsilon_{x,1} & \cdots & \epsilon_{x,N-1} \\
\epsilon_{y,0} & \epsilon_{y,1} & \cdots & \epsilon_{y,N-1} \\
0 & 0 & \cdots & 0
\end{pmatrix},
\]

where \( \epsilon_{x,i}, \epsilon_{y,i} \) have IID Gaussian distribution with zero mean and the variance \( \delta^2 \). So the perturbed trajectory matrix \( \tilde{M} = M + Z \) can be represented as:

\[
\begin{pmatrix}
x_0 + \epsilon_{x,0} & x_1 + \epsilon_{x,1} & \cdots & x_{N-1} + \epsilon_{x,N-1} \\
y_0 + \epsilon_{y,0} & y_1 + \epsilon_{y,1} & \cdots & y_{N-1} + \epsilon_{y,N-1}
\end{pmatrix}.
\]

Let us derive the perturbed null operator with the first order perturbation.

\[
\tilde{q}_0^i = -\det \begin{pmatrix} x_1 + \epsilon_{x,1} & x_2 + \epsilon_{x,2} & \cdots & x_i + \epsilon_{x,i} \\ y_1 + \epsilon_{y,1} & y_2 + \epsilon_{y,2} & \cdots & y_i + \epsilon_{y,i} \end{pmatrix}
\]

Expanding eq. (6) and ignoring the second order perturbation, it is easy to obtain:

\[
\tilde{q}_0^i = q_0^i + \epsilon_{q_0^i},
\]

where \( \epsilon_{q_0^i} \) also has Gaussian distribution with zero mean and the variance \( (x_1 - x_i)^2 + (x_2 - x_i)^2 + (x_i - x_{N-1})^2 + (y_i - y_1)^2 + (y_i - y_2)^2 + (y_1 - y_2)^2 + (y_i - y_{N-1})^2)^2 \delta^2 \), which is denoted as \( (r_{12}^2 + r_{13}^2 + r_{23}^2)^2 \delta^2 \) for simplicity. Similarly, we obtain:

\[
\tilde{q}_1^i = q_1^i + \epsilon_{q_1^i},
\]

\[
\tilde{q}_2^i = q_2^i + \epsilon_{q_2^i},
\]

\[
\tilde{q}_3^i = q_3^i + \epsilon_{q_3^i},
\]

where \( \epsilon_{q_0^i}, \epsilon_{q_1^i}, \epsilon_{q_2^i} \) satisfy:

\[
\epsilon_{q_0^i} \sim N(0, (r_{12}^2 + r_{13}^2 + r_{23}^2) \delta^2),
\]

\[
\epsilon_{q_1^i} \sim N(0, (r_{12}^2 + r_{13}^2 + r_{23}^2) \delta^2),
\]

\[
\epsilon_{q_3^i} \sim N(0, (r_{12}^2 + r_{13}^2 + r_{23}^2) \delta^2).
\]

2.2. Discussion of the error ratio

Based on the perturbed null operator, it is desirable to know the ratio of the input error and the output error where the input error is referred to the error of the trajectory matrix and the output error is referred to the error of the null operator. Let us compute the expectation of the square of
the Frobenius norm for the input error and the output error respectively. It can be shown that:

\[ E\|Z\|_F^2 = 2N\delta^2 \]  

(14)

Denoting the perturbed null operator matrix \( \tilde{Q} = \begin{pmatrix} \tilde{q}_0^0 & \ldots & \tilde{q}_{N-1}^0 \\ \tilde{q}_0^1 & \ldots & \tilde{q}_{N-1}^1 \\ \vdots & \ldots & \vdots \\ \tilde{q}_0^{N-1} & \ldots & \tilde{q}_{N-1}^{N-1} \end{pmatrix} \), we obtain:

\[ E\|Q - \tilde{Q}\|_F^2 = \sum_{i=0}^{N-1} |E(C_i^2)] + \sum_{i=0}^{N-1} (r_i^0 + r_i^1 + r_i^2) |. \]  

(15)

So the ratio of the output error and input error is:

\[ \tau = \frac{E\|Q - \tilde{Q}\|_F^2}{E\|Z\|_F^2} = \frac{1}{N} [(N - 3)(r_0^2 + r_0^2 + r_1^2) + \sum_{i=0}^{N-1} (r_i^0 + r_i^1 + r_i^2)]. \]  

(16)

From eq. (16), it can be seen that the ratio only relies on the trajectory itself while independent of the noise. Minimizing \( \tau \), we obtain:

\[ x_i = \frac{(x_0 + x_1 + x_2)}{3}, y_i = \frac{(y_0 + y_1 + y_2)}{3}. \]  

(17)

It indicates that the centroid of the first three points gives the minimum value of the ratio of the output and the input error.

2.3. Discussion of the SNR

Besides the error ratio, SNR is widely used to evaluate robustness of the system. Let us derive the expression for the output SNR. Defining the power of the output signal as \( \|Q\|_F^2 \), SNR can be computed by \( \Delta_{SNR} = \frac{\|Q\|_F^2}{E\|Q - \tilde{Q}\|_F^2} \) as:

\[ \Delta_{SNR} = \{A \sum_{i=0}^{N-1} y_i^2 + B \sum_{i=0}^{N-1} x_i^2 + C \sum_{i=0}^{N-1} x_i y_i \\
+ D \sum_{i=0}^{N-1} x_i + E \sum_{i=0}^{N-1} y_i + F \}/(28\delta^2(N - 3) \sum_{j,k=0}^{N-1} r_j^2 \\
+ \sum_{i=0}^{N-1} (r_i^0 + r_i^1 + r_i^2)) \}, \]  

(18)

where

\[ A = \sum_{j,k=0}^{2} (x_j - x_k)^2, B = \sum_{j,k=0}^{2} (y_j - y_k)^2, \]  

(19)

\[ C = -2 \sum_{j,k=0}^{2} (x_j - x_k)(y_j - y_k), \]  

(20)

\[ D = 2 \sum_{j,k=0}^{2} (y_j - y_k)(x_j y_k - x_k y_j), \]  

(21)

\[ E = 2 \sum_{j,k=0}^{2} (x_j - x_k)(x_j y_k - y_k x_j), \]  

(22)

\[ F = (N - 3) \sum_{j,k=0}^{2} (x_j y_k - x_k y_j)^2 + (x_1 y_2 - x_2 y_1 + x_0 y_1 - x_1 y_0 - x_0 y_2 + x_2 y_0)^2, \]  

(23)

where \( j \neq k \). It can be seen that the critical points of SNR are trajectory-dependent.

2.4. Bounds for the perturbation error

Using basic perturbation theorem as in [2], we obtain:

\[ \frac{\|\tilde{Q} - Q\|}{\|Q\|} \leq \|M^{-1}Z\|, \]  

(24)

where \( \|.\| \) denotes a matrix norm and a consistent vector norm. Equation (25) is referred as normwise bounds. If we weaken the bound, we can relate it with the condition number of M.

\[ \frac{\|\tilde{Q} - Q\|}{\|Q\|} \leq \|M^{-1}\| \|Z\| = \kappa(M) \frac{\|Z\|}{\|M\|}, \]  

(25)

where the condition number \( \kappa(M) = \|M\|\|M^{-1}\| \). Since the left-hand side of the bound can be regarded as a relative error in Q. The factor \( \|Z\|/\|M\| \) can likewise be regarded as a relative error in \( \tilde{M} \). Thus the condition number \( \kappa(M) \) implies how much the relative error in the matrix of the system \( MQ = 0 \) is magnified in the solution. If in addition \( \|M^{-1}Z\| < 1 \), it is easy to obtain:

\[ \frac{\|\tilde{Q} - Q\|}{\|Q\|} \leq \|M^{-1}Z\| \]  

(26)

The difficulty with normwise perturbation analysis is that it attempts to summarize a complicated situation by the relation between three numbers: the normwise relative error in \( \tilde{Q} \), the condition number \( \kappa(M) \) and the relative error in \( Q \). We can do better if we are willing to compute the inverse of M:

\[ |\tilde{Q} - Q| \leq |M^{-1}| \|Z\|\|\tilde{Q} \|. \]  

(27)
Moreover, if for some consistent matrix norm \( \|M^{-1}\|Z\| < 1 \), then \( (I - |M^{-1}\|Z|)^{-1} \) is nonnegative and
\[
|\tilde{Q} - Q| \leq (I - |M^{-1}\|Z|)^{-1}|M^{-1}\|Z||Q| . \tag{28}
\]

The bounds in (27), (28) which are referred as component-wise bounds \([4]\) can be quite an improvement over norm-wise bounds. The superiority of eqs. (15) and (18) over the bounds in the (24)-(28) are two fold: (1) Compared to (15), although the inequality (24)-(28) gives a more compact form, they are not computationally feasible when the information of the noise matrix is not available; (2) instead of inequalities, eq. (18) provides a more convenient and direct way to analyze the critical points and the convergence of SNR.

3. Sampling strategy and convergence analysis

Given the perturbation, designing optimal sampling strategy is very important for the robustness of the system. In some scenario, such as the video sequences, uniform sampling is required to ensure the quality of the video. In other cases, such as animal mobility experiments and GPS tracking, Poisson sampling is an important technique for obtaining the information.

3.1. Uniform sampling

Arbitrary trajectories in \( x \) and \( y \) directions can be represented as:
\[
\begin{align*}
  x &= f(t) \quad . \tag{29} \\
  y &= g(t) \quad . \tag{30}
\end{align*}
\]

Expanding eqs. (29) and (30) in Maclaurin series, we obtain:
\[
\begin{align*}
  f(t) &= f(0) + f' (0) t + \frac{f'' (0)}{2!} t^2 + \ldots + \frac{f^{(n)} (0)}{n!} t^n \quad . \tag{31} \\
  g(t) &= g(0) + g' (0) t + \frac{g'' (0)}{2!} t^2 + \ldots + \frac{g^{(n)} (0)}{n!} t^n . \tag{32}
\end{align*}
\]

The distance between two arbitrary samples can be computed as:
\[
\begin{align*}
  r_{kj}^2 &= (x_k - x_j)^2 + (y_k - y_j)^2 \\
  &= [f' (0) (t_j - t_k) + \ldots + \frac{f^{(n)} (0)}{n!} (t^n_j - t^n_k)]^2 + [g' (0) (t_j - t_k) + \ldots + \frac{g^{(n)} (0)}{n!} (t^n_j - t^n_k)]^2 . \tag{33}
\end{align*}
\]

With uniform sampling rate \( t_k = kT \), it is easy to obtain:
\[
\begin{align*}
  N-1 \sum_{k=3}^{N} r^2_{kj} &= \sum_{k=3}^{N-1} \left\{ \left[ \frac{f^{(n)} (0) k^n}{n!} T^n \right]^2 + \left[ \frac{g^{(n)} (0) k^n}{n!} T^n \right]^2 \right\} . \tag{34} \\
  N-1 \sum_{k=3}^{N} r^2_{k} &= \sum_{k=3}^{N-1} \left\{ \left[ \frac{f^{(n)} (0)}{n!} (k-1)^T + \ldots + \frac{f^{(n)} (0)}{n!} \right]^2 + \left[ \frac{g^{(n)} (0)}{n!} (k-1)^T + \ldots + \frac{g^{(n)} (0)}{n!} \right]^2 \right\} . \tag{35} \\
  N-1 \sum_{k=3}^{N} r^2_{2k} &= \sum_{k=3}^{N-1} \left\{ \left[ \frac{f^{(n)} (0) k^n}{n!} T^n \right]^2 + \left[ \frac{g^{(n)} (0) k^n}{n!} T^n \right]^2 \right\} . \tag{36}
\end{align*}
\]

Therefore, it can be seen from eqs. (34)-(36) that \( \tau \) is equal to \( O(N^{2n}) \) which increases dramatically with the number of the samples. With regard to SNR, we have the following property:

**Property 1:** With uniform sampling \( t_k = kT \),
\[
\lim_{N \to \infty} \Delta_{SNR} = \frac{A (g^{(n)} (0))^2 + B (f^{(n)} (0))^2 + C f^{(n)} (0) g^{(n)} (0)}{6 \delta^2 [(f^{(n)} (0))^2 + (g^{(n)} (0))^2]} . \tag{37}
\]

where \( A, B \) and \( C \) are defined in (19) and (20).

**Remark:** Property 1 can be proved by showing that both \( \|Q\|^2 \) and \( E\|Q - \tilde{Q}\|^2 \) are \( O(N^{2n+1}) \). Thus, SNR is determined mainly by three factors: (1) The coordinates of the first three samples; (2) the values of the \( n \)th derivatives at the origin; and (3) the variance of the noise.

3.2. Poisson sampling

To guarantee the convergence of the error ratio, Poisson sampling is chosen which has the distribution for the sampling time \( t_k \) as:
\[
\begin{align*}
  \mathcal{P}(t_k) &= \frac{\lambda^{k} k^{-1} e^{-\lambda t}}{(k-1)!} . \tag{38}
\end{align*}
\]

The \( n \)th moment of \( t_k \) can be expressed as:
\[
\begin{align*}
  E(t^n_k) &= \int_{0}^{\infty} \frac{\lambda^{k} n + k - 1 e^{-\lambda t}}{(k-1)!} dt \\
  &= \frac{(n + k - 1)!}{\lambda^{n}(k-1)!} . \tag{39}
\end{align*}
\]

**Property 2:** \( \lambda = O(N) \) should be chosen for Poisson sampling to guarantee the convergence of \( \tau \), where \( N \) is the total


number of samples.

**Property 3:** \( \tau \) converges in mean sense given \( \lambda = O(N) \). Specifically, for \( \lambda = \frac{N}{\sqrt{N}} \),

\[
\lim_{N \to \infty} E(\tau) = 3 \sum_{k=0}^{2n} \frac{(a_k + b_k)T^k}{k + 1}, \tag{40}
\]

where \( k \) is the index for Taylor series and for \( a_k \) and \( b_k \), if \( k \) is odd,

\[
a_k = \sum_{i=1}^{\frac{k-1}{2}} \frac{2g^{(i)}(0)g^{(k-i)}(0)}{i!(k-i)!}, \tag{41}
\]

\[
b_k = \sum_{i=1}^{\frac{k-1}{2}} \frac{2f^{(i)}(0)f^{(k-i)}(0)}{i!(k-i)!}, \tag{42}
\]

if \( k \) is even,

\[
a_k = \sum_{i=1}^{\frac{k}{2}} \frac{2g^{(i)}(0)g^{(k-i)}(0)}{i!(k-i)!} + \frac{(g^{(k)}(0))^2}{(k!)^2}, \tag{43}
\]

\[
b_k = \sum_{i=1}^{\frac{k}{2}} \frac{2f^{(i)}(0)f^{(k-i)}(0)}{i!(k-i)!} + \frac{(f^{(k)}(0))^2}{(k!)^2}. \tag{44}
\]

**Property 4:** If the trajectories are sampled with \( \lambda = O(N) \), the variance of the error ratio converges to zero, namely,

\[
\lim_{N \to \infty} Var(\tau) = 0. \tag{45}
\]

**Remark:** Property 3 can be proved by showing that

\[
\lim_{N \to +\infty} \frac{1}{N} \sum_{j=1}^{N-1} E(t_j^k) = \frac{1}{k + 1}. \tag{46}
\]

In our framework, the density \( \lambda \) corresponds to the average number of samples per unit-length; i.e. \( \lambda = \frac{N}{\sqrt{N}} \).

### 4. Simulation Results

In this section, we first demonstrate the superiority of the performance by Poisson sampling over uniform sampling in terms of the error ratio for different trajectories. Then we show the performance of the SNR with uniform sampling outperforms the one with Poisson sampling. Moreover, we evaluate the performance of our system for retrieval and classification of motion trajectories with both perfect data and noisy data relying on principal component null space analysis (PCNSA) from the Context Aware Vision using Image-based Active Recognition (CAVIAR) data set in [1].

#### 4.1. Evaluation of the error ratio and SNR

For numerical examples, we extract the true trajectory from Fig 1 and evaluate the performance of \( \tau \) over \( N = 60 \) samples with \( \lambda = N, N/2, N/3 \) respectively. Shown in Fig.2, for the three cases, Poisson sampling works well and among them, the case of \( \lambda = N \) gives the best performance. Let’s define \( SNR(dB) = 10 \log_{10}(\Delta_{SNR}) \). The perturbed trajectories are chosen with zero mean and the variance 0.5 additive Gaussian noise to evaluate the performance of SNR by uniform sampling over 30 samples with the sampling period \( T=1s \) from Fig.1. It can be seen from the simulation results in Fig.3 that (1) the general trend is that SNR decreases when the number of samples \( N \) increases; (2) the convergence of SNR is very fast. Theoretically, there is no closed form for the mean of the SNR with Poisson sampling. We also provide the comparison of the performance of SNR with uniform sampling and Poisson sampling for the same trajectories we extracted in Fig.3. In Fig.3 the performance of SNR with uniform sampling is superior to the performances with Poisson sampling for \( \lambda = N/4, N/5 \) respectively.

#### 4.2. Trajectory retrieval and classification

Since in the real world, the trajectories in a class usually have different lengths, we normalize the length by taking Fourier Transform and choosing the largest 32 coefficients and then taking Inverse Fourier Transform so that all the trajectories are of size 32. Based on NSI, the \( N \times (N - 3) \) matrix \( Q \) is converted into \( n(n-3) \) column vector \( Y \) which is considered as data samples. We perform Principal Component Null Space Analysis (PCNSA) with the discriminant function \( D(X_i, Y) = ||W_{NSA,i}(X_i - Z)|| \), where \( Z \) is the query trajectory. The PR curve for motion trajectory retrieval is shown in Fig.4. The visual illustration of the query and the three most similar retrieval are shown in Fig.5. In Fig.6, we use 20 different classes as different types of motions in the data set. Each class has 50 trajectories recorded at different instances. Based on PCNSA, We plot the PR curve for the three cases, Poisson sampling works well and among them, the case of \( \lambda = N \) gives the best performance.
Figure 3. SNR(dB) comparison vs the number of samples with Poisson sampling $\lambda = N/4, N/5$ and uniform sampling for the motion trajectories in Fig.1.

Figure 4. Precision-Recall Curve for indexing and retrieval with 20 classes with true and noisy trajectories.

Figure 5. Visual illustration for retrieval results with 20 classes with motion trajectories.

Figure 6. Precision-Recall Curve for classification with 40 classes with true and noisy trajectories.

5. Conclusion

In this paper, we propose a novel robust system to apply the null space invariants to computer vision and other related fields. 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 preserves information of the original data when computing invariants. We derive the perturbed null operators with the first order perturbation theory and analyze the sensitivity of the systems in terms of the error ratio and the SNR. In view of the perturbation, optimal sampling are purposed to minimize the corresponding sensitivity. Computer simulation demonstrates the effectiveness and robustness of our approach in indexing, retrieval and classification of both perfect and noisy motion trajectories.

References