SAND2018-7455C

Koopman Mode Decomposition of Pairwise Distance Observables from Feature Points

Cory N. Brown[1] Ryan Mohr[1,3] Igor Mezić[1,2,3]
[1]Department of Mechanical Engineering, and
[2]Department of Mathematics,
University of California, Santa Barbara, CA, 93106.
[3]AIMdyn, Inc., Santa Barbara, CA, 93106.
corybrown@ucsb.edu
mezic@ucsb.edu

Akshat Kumar
Sandia National Laboratories, Livermore, CA 94550.

Abstract

This paper explores the problem of motion segmentation in video sequences in the context of multibody motion tracking with a moving camera. A common approach is to use the locations of tracked feature points and cluster their trajectories into different motions. Some motion tracking algorithms will also include object detection. Herein we consider the use of a relatively recent framework to view dynamical systems theory, which naturally finds application in data science, known as Koopman spectral analysis (KSA). In particular, we apply KSA to times series of distances between feature points. We do this through the use of dynamic mode decomposition (DMD) with pairwise distances of feature points as observables. DMD is a relatively inexpensive algorithm that can extract spatio-temporal patterns which have intuitive dynamic interpretations. We demonstrate the use of DMD on data from tracked feature points in multiple video sequences without the use of any object recognition. We show that DMD can be used to generate features which could be exploited by other machine learning tools.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

1. Introduction

Autonomous cars, video surveillance, multi-robot collaboration, and medical imaging are all applications which involve analysis and classification of large amounts of dynamic image sequences in real time. Better approaches to these problems have largely been obtained by addressing smaller ones of computer vision, e.g. object detection, identification, tracking, motion estimation in 3D, and forecasting of future dynamics, or integrating their advancements. The idea of *motion segmentation*, the partitioning of images, or trajectories of tracked feature points, in a video into regions of similar dynamics so that each element of the partition has a corresponding element in the partition of the previ-

ous or following image, plays a key role in this area of research.

Object tracking has mostly been approached through unsupervised methods [23, 29, 26]. One might think that the problem would be greatly simplified by using object detection. However, using purely bottom up approaches in combination with detection seems to be optimal [1, 7, 11]. Unsupervised approaches are surely justified in contexts where tracking undetectable objects is necessary. In addition, they can be used to make object detection more robust and computational efficient and streamline the training of object detectors by finding undetectable moving objects in large amounts of video sequences.

In the early 1930s, Bernard Koopman and John von Neumann introduced a new operator theortic view of dynamical systems [14, 13]. In the 2000s, it was discovered to be a useful framework for data-driven analysis of high-dimensional nonlinear systems [22, 20]. Since then, numerical methods used to approximate Koopman spectral objects have been used as data analysis tools in the contexts of fluid dynamics [2], neurodynamics [5], energy efficiency [8], and molecular dynamics [31]. A particular class of algorithms for approximating Koopman spectral objects, known as dynamic mode decomposition [27, 25], has even been used in background/foreground separation in videos taken from a stationary camera [17]. Herein, we consider videos taken from a moving camera and perform dynamic mode decompostion on a different set of *observables* - measurements we take of a dynamical system (defined more precisely in the next section).

Standard observables used in the context of motion segmentation included feature point coordinates [24] and optical flow[10] components. In combintation with DMD, Kutz et al. used grayscale pixel data [17]. One element of the work herein which distinguishes it is the choice of observables - pairwise distance of feature points.

The paper is organized as follows: In section 2 we give a brief introduction to the Koopman view of dynamical systems. In section 3 we introduce dynamic mode decomposition (DMD) and explain its relation to Koopman spectral analysis. In section 4 we demonstrate the use of Koopman spectral analysis on pairwise distance observables of feature points for several movies through the use of DMD. We conclude in section 5.

2. The Koopman View of Dynamical Systems

In the context of deterministic dynamical systems, state-variables are the variables whose knowledge at a particular time determines their values for all of time (e.g. angle and angular velocity for the mathematical pendulum). The classical approach in dynamical systems theory is to work in a mathematical framework centered around the state-space viewpoint. When recording dynamic data from complex systems such as video sequences, the state variables which fully describe the observed system are not known and there is no model to describe their evolution. In these circumstances, the Koopman view of dynamical systems, in which the main objects of interest are functions whose domain is the state-space (called *observables*), is a natural one to take. In the case of the pendulum, observables include kinetic energy, potential energy, or any function of angle and angular momentum one could think of. In this view, we can think of any measurements we take of a system of interest as a function of some unknown state. In addition, we can always construct new observables from measured ones - e.g. composing vectors of measured observables on the left with any complex valued function on n-dimensional complex space or on the right with any function from the state-space to itself. In the context of video sequences, observables then include pixel data, optical flow field components at each location, and feature point locations.

The simplest context in which to introduce the Koopman view is for discrete-time dynamical systems which can be described by repeated application of a single function $T : M \to M$, i.e.

$$x' = T(x) \tag{1}$$

where M is the state-space and x' is the updated state, corresponding to x, after a single time step.

The reader should note that the Koopman framework was originally described for time-invariant continuous-time systems [14] and has been extended to the case of time-varying[19] and even random and stochastic[6] dynamical systems. For the dynamical system described by Eq. 1 the induced Koopman operator U represents a new discrete-time dynamical system whose state-space is the set of all complex-valued functions with domain M (denoted by $\mathbb{C}^M$):

$$f' := U(f) = f \circ T. \tag{2}$$

In Eq. 2, the symbol $\circ$ denotes function composition. It is easy to show that U is linear ($c \in \mathbb{C}$ and $f, g \in \mathbb{C}^M$),

$$\begin{aligned} U(cf+g) &= (cf+g) \circ T && (3)\\ &= c(f \circ T) + g \circ T && (4)\\ &= cU(f) + U(g), && (5) \end{aligned}$$

thus we can consider eigenvalues λ and eigenfunctions ϕ,

$$U(\phi) = \lambda\phi. \tag{6}$$

In particular, we call λ and ϕ Koopman eigenvalues and Koopman eigenfunctions, respectively, of T.

Let $\{f_1, \ldots, f_n\}$ denote a finite collection of observables on the state-space of T. Assume that $\{f_1, \ldots, f_n\}$ lie in the span of m eigenfunctions $\{\phi_1, \ldots, \phi_m\}$ of U associated with eigenvalues $\{\lambda_1, \ldots \lambda_m\}$. Let $\{\psi_1, \ldots, \psi_n\}$ denote the dual basis of $\{\phi_1, \ldots, \phi_m\}$. We can then write a special case of the Koopman Mode Decomposition (KMD) for the q times repeated action of U on a finite collection of n observables $\{f_1, \ldots, f_n\}$:

$$\begin{aligned} \begin{bmatrix} U^q(f_1) \\ \vdots \\ U^q(f_n) \end{bmatrix} &= \begin{bmatrix} \sum_{k=1}^{m} \lambda_k^q \phi_k \psi_k(f_1) \\ \vdots \\ \sum_{k=1}^{m} \lambda_k^q \phi_k \psi_k(f_n) \end{bmatrix} && (7)\\ &= \sum_{k=1}^{m} \lambda_k^q \phi_k \begin{bmatrix} \psi_k(f_1) \\ \vdots \\ \psi_k(f_n) \end{bmatrix}. && (8) \end{aligned}$$

This is simply an eigendecomposition in each component acting on different vectors in each component where the $\psi_k s$ are linear coordinate functions. This decomposition allows us to evolve our observables $\{f_1, \ldots, f_n\}$ by simply multiplying each term in the sum above by its corresponding eigenvalue λ_k. This gives us the intuitve interpretations of the magnitude and complex phase of λ_k as corresponding to rate of growth and rate of oscillation, respectivley. If we are only interested in the evolution of the observables along a single trajectory of (1) starting at x, then the ϕ_ks could be chosen so that $\phi_k(x) = 1$, and KMD would take the following simpler form

$$\begin{bmatrix} [U^p(f_1)](x) \\ \vdots \\ [U^p(f_n)](x) \end{bmatrix} = \sum_{k=1}^{m} \lambda_k^p \begin{bmatrix} \psi_k(f_1) \\ \vdots \\ \psi_k(f_n) \end{bmatrix}. \tag{9}$$

The last column vector appearing in (8) and (9) is known as the Koopman mode corresponding to the pair (λ_k, ϕ_k) and the vector of observables $[f_1, \ldots, f_n]^T$ relative to the discrete-time system descibed by (1). From (9) we see that the magnitude of $\psi_k(f_i)$ tells us how much the growth and oscillations rates (as a pair) given by λ_k play a role in the evolution of f_i along the single trajectory starting at x and locally in time. Similarly, the complex phase of $\psi_k(f_i)$ gives a relative phase correspsonding to the oscillations give by λ_k.

As an example, consider the special case that M is a n-dimensional Euclidean space and T is a diagonalizable linear operator. In addition, let $\{f_1, \ldots, f_n\}$ be the dual basis of the standard ordered basis for $\mathbb{C}^n$. Given any basis $\{v_1, \ldots, v_n\}$ of eigenvectors for T with eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$, we have that its dual basis $\{\phi_1, \ldots, \phi_n\}$ is a set of eigenfunctions for U with eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$. We see that for any

$x \in M = \mathbb{R}^n$, we have

$$\begin{bmatrix} [U^p(f_1)](x) \\ \vdots \\ [U^p(f_n)](x) \end{bmatrix} = \begin{bmatrix} f_1(T^p(x)) \\ \vdots \\ f_n(T^p(x)) \end{bmatrix} \tag{10}$$

$$= T^p(x) \tag{11}$$

$$= \sum_{k=1}^{m} \lambda_k^p \phi_k(x) v_k. \tag{12}$$

By comparing with equation (8) we see that $\{v_1, \ldots, v_n\}$ are the Koopman modes corresponding to the pairs $\{(\lambda_1, \phi_1), \ldots, (\lambda_n, \phi_n)\}$ and the vector of observables $[f_1, \ldots, f_n]^T$, relative to the discrete-time system descibed by (1). With this example we see that eigendecomposition can be viewed as a special case of KMD; the extension that KMD provides is flexibilty in both $\{f_1, \ldots, f_n\}$ and T, in particular, these can be nonlinear.

For more detail on the spectral expansion of the Koopman operator for data-driven analysis of the underlying dynamical system see [22, 20].

3. Koopman Spectral Analysis and Dynamic Mode Decomposition

Consider the case where we have a discrete-time dynamical system as in (1) and we have evaluated a set of observables $\{f_1, \ldots, f_n\}$ along the first $m+1$ time points of a trajectory starting at x. We can put this into a matrix D such that each row corresponds to a different observable and the columns are ordered by time:

$$D = \begin{bmatrix} f_1(x) & f_1(T(x)) & \ldots & f_1(T^m(x)) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(x) & f_n(T(x)) & \ldots & f_n(T^m(x)) \end{bmatrix} \tag{13}$$

$$= \begin{bmatrix} f_1(x) & [U(f_1)](x) & \ldots & [U^m(f_1)](x) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(x) & [U(f_n)](x) & \ldots & [U^m(f_n)](x) \end{bmatrix}. \tag{14}$$

We can split D into the matrix X of the first m columns and Y of the last m columns, i.e.

$$X = \begin{bmatrix} f_1(x) & f_1(T(x)) & \ldots & f_1(T^{m-1}(x)) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(x) & f_n(T(x)) & \ldots & f_n(T^{m-1}(x)) \end{bmatrix}, \tag{15}$$

$$Y = \begin{bmatrix} f_1(T(x)) & f_1(T^2(x)) & \ldots & f_1(T^m(x)) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(T(x)) & f_n(T^2(x)) & \ldots & f_n(T^m(x)) \end{bmatrix}. \tag{16}$$

The main idea in dynamic mode decomposition (DMD) is to find a matrix A such that AX is close to Y in some sense. In this way A would be mapping the vector of measurements taken at the point $T^i(x)$ close to those taken at the point $T^{i+1}(x)$, for all $i \in \{0, \ldots, m-1\}$. With this, it may seem intuitive to the reader that as our number of time points goes to infinity or as we add more observables (which could just be functions of our original ones), we should better and better approximate the Koopman operator U (and its spetral objects) by the linear operator we are representing by the matrix A; in fact, in some cases this has be proven to be the case [30, 12, 16, 3, 21]. One such A we could use is $A = YX^\dagger$, where $X^\dagger$ is the pseudo-inverse of X. Such an A happens to satsfy the following [4]

$$\|AX - Y\|_F = \inf_{B \in \mathbb{R}^{n \times n}} \|BX - Y\|_F, \tag{17}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Note that we are particularly considering extended dyanmic mode decomposition. [16, 30]

Consider $S : \mathbb{R}^n \to \mathbb{R}^n$ defined by $S(y) = Ay$ - known as the DMD operator. Since this is a linear operator on $\mathbb{R}^n$, a Koopman mode decompostion of the dual basis to the standard order basis on $\mathbb{R}^n$, with respect to the discrete time map

$$y' = S(y), \tag{18}$$

is given by an eigendecomposition of S as explained at the end of the previous section. It is common to call these resulting Koopman eigenvalues, eigenfunctions, and modes (coming from the DMD operater), DMD eigenvalues, eigenfunctions, and modes, respectively.

4. Koopman Mode Decomposition of Pairwise Distance Observables

In this section we apply Koopman spectral analysis to three real-life videos (with a moving camera) via the use of DMD. The first image of each sequence can be found in the first row of figure 1. From left to right the image sequences increase in their number of motions from two to four where the background is treated as a motion since the camera is moving. Note that feature points belonging to the same motion have the same color and feature point trajectories for the length of the video are plotted in the second row. The first video was taken from the data set used in [28, 29]. The second video was shared by the authors of [23]. Finally, the last video is from the data set used in [24]. With each of these movies we have a set of feature points $\{p_1, \ldots, p_r\}$, where we make the identification

$$p_i = (x_{i1}, \ldots, x_{im}, y_{i1}, \ldots, y_{im}) \tag{19}$$

with x_{ij} and y_{ij} denoting the x and y coordinates of feature point i in frame j. Thus, we are letting r and m denote the number of feature points and the number of frames, respectively. In the four motion sequence, some of the feature points are outliers[24] (see figure 1).

The main assumption here is that there is some underlying dynamical system which describes the real-world motions which were captured by a video camera, with a constant frame rate, and that we can represent the restriction of this dynamical system to a discrete time set (with difference between time points corresponding to the frame rate) as in (1). We can then use DMD by choosing a set of observables. For a single movie, we view each color or grayscale component of each pixel as providing us with an observable along a single trajectory; even though we do not know the state-space of the underlying system we are observing,i.e. the domain of T, we assume that the processes of capturing an image is a sampling of a set of observables at a single point in state space and at a single point in time. The coordinates of a feature point can then be thought of as real-valued functions composed with a vector of our pixel level observables. In this way, the coordinates of each feature point are again observables. Finally, pairwise Euclidean distances of feature points are the observables which will be subject to our Koopman spectral analysis.

We denote the value of the Euclidean distance between feature point i and j in frame k by

$$d_{ijk} = ||(x_{ik}, y_{ik}) - (x_{jk}, y_{jk})||_2. \tag{20}$$

Using these values, except those pairwise distances with $j \leq i$ (to keep prevent unnecessary redundancy), we construct the vectors

$$D_k = [d_{12k} \cdots d_{1rk}\ d_{23k} \cdots d_{2rk}\ d_{(r-1)rk}]^T. \tag{21}$$

Thus D_k is a vector of length $1 + \cdots + (n-1) = \frac{n(n-1)}{2}$ and we can use the matrices

$$X = \begin{bmatrix} | & | & & | \\ D_1 & D_2 & \cdots & D_{(m-1)} \\ | & | & & | \end{bmatrix}, \text{ and} \tag{22}$$

$$Y = \begin{bmatrix} | & | & & | \\ D_2 & D_3 & \cdots & D_m \\ | & | & & | \end{bmatrix}. \tag{23}$$

in a DMD algorithm. In particular we used the DMD algorthim in [5], which includes a way to assign a "power" that each DMD mode represents in the data matrix X. The resulting power is higher if the mode is more parallel to left singular vectors with large corresponding singular values and low otherwise. We can use this to order modes based on the level of important they play in our data matrix as opposed to using the coefficients from the projection of a column of the data matrix onto the DMD modes. Finally, since the algorithm invovles inverses of the diagonal matrix of singular values of X, we treat any singular value less than $\frac{\sigma}{10^6}$ as zero, where σ is the largest singular value.

Using the DMD algorithm mentioned, we obtain DMD eigenvalues and modes. The DMD eigenvalues are plotted in row three of figure 1 and are colored according to their associated percent total power. The DMD modes corresponding to the the eigenvalues with the top three highest powers (identifying complex conjugate pairs), are represented in the remaining three rows of figure 1. We

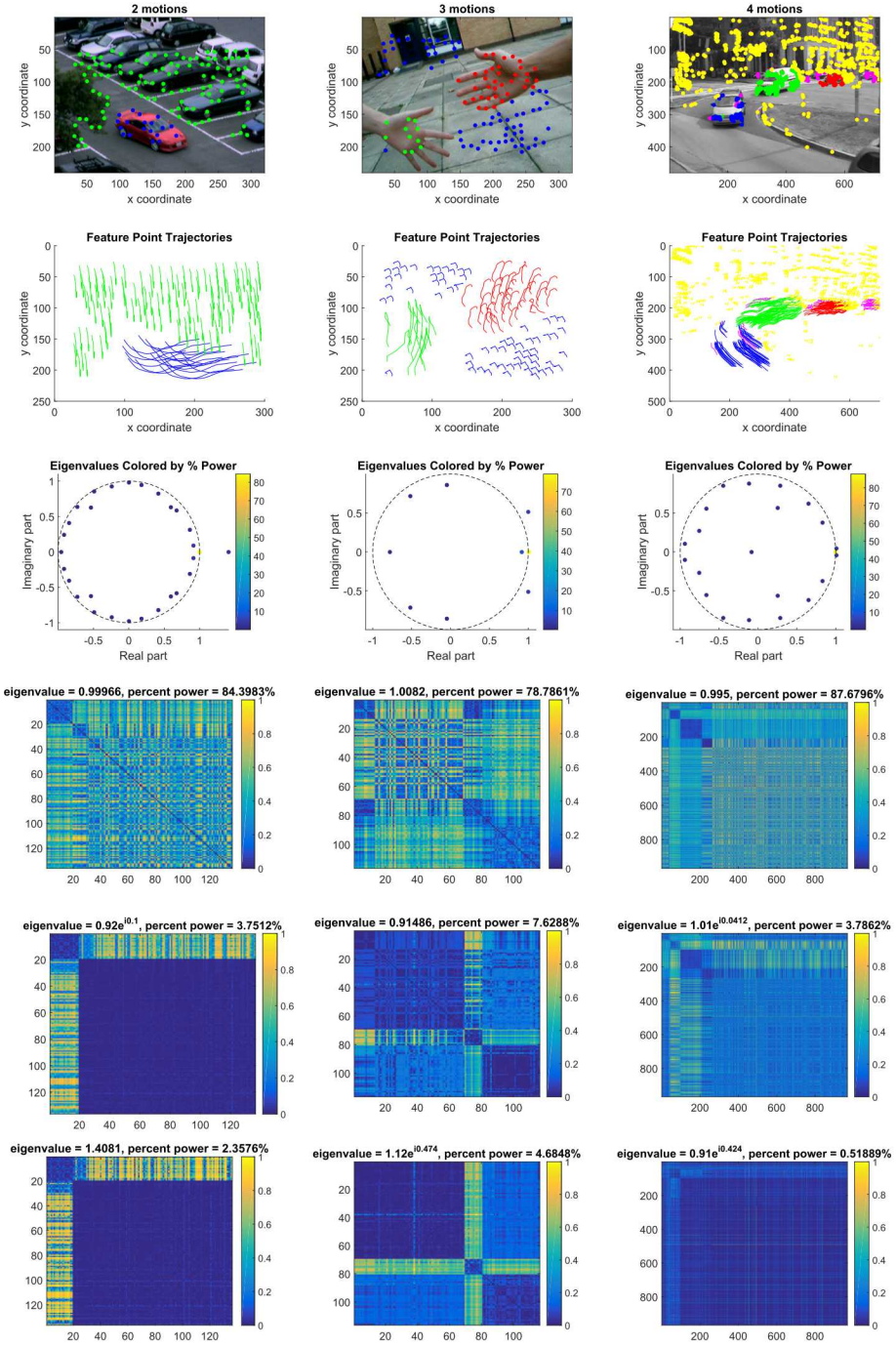

Figure 1: Caption on next page.

Figure 1: Koopman spectral quantities of the DMD matrix obtained from pairwise distance observables on three real-world video sequences. Each movie and its computed Koopman quantities correspond to a single column. The first row contains a picture from each video sequence with feature points labeled by different motions or outliers. In the second row, the paths the feature points trace are plotted, for the full video sequence, and colored to match the pictures above them. The third row contains eigenvalues of the DMD matrix colored by their percent power. The last three rows contain the three highest power eigenvectors of the DMD matrix represented as symmetric matrices; the i, j-th entry corresponds to the pairwise distance observable between the ith and jth feature points. The rows (and columns) of these matrices are ordered by motion with outliers first (if any). Note that in the 4 motion movie, the last 15 frames were not used due to obvious mistracked feature points in those frames.

represent the modes by symmetric matrices such that the $i, j-$th element corresponds to the pairwise distance between feature points i and j. Since these modes are in general complex, we only plot the magnitude of the components. Our feature points are already labeled according to different motions, thus we order the rows and columns of these matrices by color of the corresponding feature point: magenta (outiers), blue, green, red, and, finally, yellow. This leads to some block structure which is evident from the figure.

The first thing to notice is that the highest power DMD mode and eigenvalue pair has the eigenvalue closest to one in all three examples. We call this the stationary mode since it evolves in time by repeatedly multiplying it by one. For the two motion video (first column), this mode (fourth row) has a 19 by 19 principle submatrix in the upper left corner which is almost zero. We also see this in the lower power modes. This block appears in the stationary mode because the feature points are close together and in the remaining (dynamic) modes because these feature points do not move much relative to eachother. On the other hand, the remaining motion of the two motion video is the background; the background points are fairly spreadout and so we do not see a similar block of zeros in the stationary mode as we did for the localized object (car). However, we do see such blocks in the dynamic modes. As one might expect, the dynamic modes have large components outside of the zero blocks on the diagonal; this is due to the relative motion of the car and the background. These symmetric matrix representations of the dynamic modes are static pictures which capture dynamic information.

Looking to the second column of figure 1, we notice similar structure in the DMD modes as mentioned for the two motion video. All three plotted modes have two fairly homogeneous blue blocks on the diagonal, in the bottom right corners, which correspond to the green hand (smaller block) and red hand (bottom right). A similar block for the backgound points only appears in the dynamic modes. As with the two motion case, the dynamic modes have larger components outside of these zero blocks on the diagonal. Similar structures are again seen in modes for the remaining four motion video (third column of figure 1).

5. Conclusion and Outlook

The main contribution of this work, distinguishing it from that in [17], is the consideration of KMD on pairwise distance observables of tracked feature points as opposed to grayscale intensity observables. The coordinates of these feature points, derived from optical flow based video processing methods such as the Kanade-Lucas-Tomasi feature tracking algorithm, yield a sparser spatial dataset for processing than the full pixel-space. Moreover, using pairwise distance observables derived from these coordinates not only gives us many more meaningful observables, it also leads to the intuitive block structures seen in figure 1. In addition, the intensity-based algorithm of [17] relies on separation of slow versus fast time-scales for separating motions in different spatial regions with similar time-scales; in contrast, the pairwise distance observables separate local motions at varying speeds and directions in a single instance of the algorithm and are independent of the color/intensity space.

Finally, while the results in [17] were impressive and multi-resolution DMD has many applications other than object tracking, the approach does not seem to easily extend to videos with a dynamic background. Figure 1 above shows that Koopman modes of pairwise distance observables can segment motion even with a moving background.

We have represented the absolute value of DMD modes obtained from pairwise distance observables as non-negative symmetric matrices which can be interpreted as a distance matrix for a graph whose nodes are the tracked feature points. An amazing feature of this approach is that object separation can be obtained from a single dynamic mode; this is very clearly seen, for example, in the bottom left plot of the figure. Similar to how the problem of motion segmentation is stated in [24], one approach to clustering would be to permute these matrix representations, and obtain the approximately block-diagonal structures as seen in figure 1. On the other hand, [23] achieves image segmentation by clustering data by eigenvectors of a certain distance (commute time) matrix and suggest this could also be done with feature point data from videos. Thus another possible route to automatically separating the motions based on DMD modes would be to cluster based on the eigenvectors of their non-negative symmetric matrix representations.

From the above discussion it is clear that choice of observables is key to the success of applying KMD to data. We merely note that observables can be choosen using a neural network approach [32, 18].

In applications involving non-rigid motions, such as tracking people, their may me significant relative motion of feature points on a single object to be tracked. In the case of tracking a few people walking, this may result in serveral dynamic modes, represented as in figure 1, with a single nonzero block on the diagonal and zeros elsewhere; we would expect such modes to have corresponding eigenvalues with oscillation rates given by the different cadences of the walkers.

In applications such as autonomous cars, estimating and forecasting motions of tracked objects is also relevant. We note that KMD can be used for prediction of future dynamics [9, 15] and coordinates of estimated motions of tracked objects would be the natural observables to use.

References

[1] M. Andriluka, S. Roth, and B. Schiele. People-tracking-by-detection and people-detection-by-tracking. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[2] H. Arbabi, M. Korda, and I. Mezic. A data-driven koopman model predictive control framework for nonlinear flows. *arXiv preprint arXiv:1804.05291*, 2018.

[3] H. Arbabi and I. Mezic. Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the koopman operator. *SIAM Journal on Applied Dynamical Systems*, 16(4):2096–2126, 2017.

[4] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[5] B. W. Brunton, L. A. Johnson, J. G. Ojemann, and J. N. Kutz. Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of neuroscience methods*, 258:1–15, 2016.

[6] N. Črnjarić-Žic, S. Maćešić, and I. Mezić. Koopman operator spectrum for random dynamical system. *arXiv preprint arXiv:1711.03146*, 2017.

[7] B. Drayer and T. Brox. Object detection, tracking, and motion segmentation for object-level video segmentation. *arXiv preprint arXiv:1608.03066*, 2016.

[8] M. Georgescu and I. Mezić. Building energy modeling: A systematic approach to zoning and model reduction using koopman mode analysis. *Energy and buildings*, 86:794–802, 2015.

[9] D. Giannakis. Data-driven spectral decomposition and forecasting of ergodic dynamical systems. *Applied and Computational Harmonic Analysis*, 2017.

[10] B. K. Horn and B. G. Schunck. Determining optical flow. *Artificial intelligence*, 17(1-3):185–203, 1981.

[11] Z. Kalal, K. Mikolajczyk, and J. Matas. Tracking-learning-detection. *IEEE transactions on pattern analysis and machine intelligence*, 34(7):1409–1422, 2012.

[12] S. Klus, P. Koltai, and C. Schütte. On the numerical approximation of the perron-frobenius and koop-

man operator. *arXiv preprint arXiv:1512.05997*, 2015.

[13] B. Koopman and J. v. Neumann. Dynamical systems of continuous spectra. *Proceedings of the National Academy of Sciences*, 18(3):255–263, 1932.

[14] B. O. Koopman. Hamiltonian systems and transformation in hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931.

[15] M. Korda and I. Mezić. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *arXiv preprint arXiv:1611.03537*, 2016.

[16] M. Korda and I. Mezić. On convergence of extended dynamic mode decomposition to the koopman operator. *Journal of Nonlinear Science*, 28(2):687–710, 2018.

[17] J. N. Kutz, X. Fu, S. L. Brunton, and N. B. Erichson. Multi-resolution dynamic mode decomposition for foreground/background separation and object tracking. In *Computer Vision Workshop (ICCVW), 2015 IEEE International Conference on*, pages 921–929. IEEE, 2015.

[18] Q. Li, F. Dietrich, E. M. Bollt, and I. G. Kevrekidis. Extended dynamic mode decomposition with dictionary learning: A data-driven adaptive spectral decomposition of the koopman operator. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(10):103111, 2017.

[19] S. Maćešić, N. Črnjarić-Žic, and I. Mezić. Koopman operator family spectrum for nonautonomous systems-part 1. *arXiv preprint arXiv:1703.07324*, 2017.

[20] I. Mezić. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, 41(1-3):309–325, 2005.

[21] I. Mezic and H. Arbabi. On the computation of isostables, isochrons and other spectral objects of the koopman operator using the dynamic mode decomposition.

[22] I. Mezić and A. Banaszuk. Comparison of systems with complex behavior. *Physica D: Nonlinear Phenomena*, 197(1-2):101–133, 2004.

[23] H. Qiu and E. R. Hancock. Clustering and embedding using commute times. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11), 2007.

[24] S. Rao, R. Tron, R. Vidal, and Y. Ma. Motion segmentation in the presence of outlying, incomplete, or corrupted trajectories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(10):1832–1845, 2010.

[25] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson. Spectral analysis of nonlinear flows. *Journal of fluid mechanics*, 641:115–127, 2009.

[26] R. Sabzevari and D. Scaramuzza. Multi-body motion estimation from monocular vehicle-mounted cameras. *IEEE Transactions on Robotics*, 32(3):638–651, 2016.

[27] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics*, 656:5–28, 2010.

[28] Y. Sugaya and K. Kanatani. Outlier removal for motion tracking by subspace separation. *IEICE TRANSACTIONS on Information and Systems*, 86(6):1095–1102, 2003.

[29] Y. Sugaya and K. Kanatani. Multi-stage unsupervised learning for multi-body motion segmentation. *IEICE Transactions on Information and Systems*, 87(7):1935–1942, 2004.

[30] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley. A data–driven approximation of the koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015.

[31] H. Wu, F. Nüske, F. Paul, S. Klus, P. Koltai, and F. Noé. Variational koopman models: slow collective variables and molecular kinetics from short off-equilibrium simulations. *The Journal of Chemical Physics*, 146(15):154104, 2017.

[32] E. Yeung, S. Kundu, and N. Hodas. Learning deep neural network representations for koopman operators of nonlinear dynamical systems. *arXiv preprint arXiv:1708.06850*, 2017.