Coding Project 5: Background Subtraction through Dynamic Mode Decomposition

Manjaree Binjolkar

Abstract

The aim of the study was to use Dynamic Mode Decomposition (DMD) to subtract the background variations in order to analyze the temporal dynamics of a moving object in the video clip and extract its spatiotemporal patterns.

1 Introduction

Dynamic Mode Decomposition (DMD) is a data-driven method used to identify coherent structures and dynamic behaviors within time-varying systems. The technique works by decomposing the system's high-dimensional snapshots into a set of spatial modes and their corresponding temporal dynamics, which can be used to reconstruct and predict the system's behavior. DMD is particularly useful when traditional analytical methods are not feasible due to the complexity of the system or the lack of a mathematical model. DMD has applications in a variety of fields, including fluid dynamics, neuroscience, image processing, and climate modeling, among others.

One significant advantage of DMD is that it does not require modifying the system or collecting additional data. It is based solely on data collected from the system, it can provide insight into the underlying physics or mechanisms driving the system's behavior which makes it an attractive tool for researchers and engineers. This information can be used to improve system performance, optimize design, or even develop control strategies.

In this report, we used DMD on the video clip phi4.mov and attempted to separate its foreground and background variations.

2 Theoretical Background

The mathematical formulation of DMD involves the Singular Value Decomposition of a data matrix X, which is constructed by stacking the columns of the input video stream as vectors. The DMD modes can be obtained from the dominant singular vectors and values, which capture the spatiotemporal patterns of motion in the video stream. It takes advantage of low dimensionality in the experimental data itself without having to rely on a given set of governing equations. It can be viewed as computing the eigenvalues and eigenvectors (low-dimensional modes) of a linear model from the experimental data in order to approximate the underlying dynamics, even though the dynamics are nonlinear.

Assume we have a set of time-varying data points $X = x_1, x_2, ..., x_n$, where x_i is a vector in $\mathbb{R}^d$ representing the state of the system at time i. The goal of DMD is to find a set of linear equations that describe the evolution of the system over time. Specifically, we want to find a matrix A such that $Ax_i \approx x_{i+1}$ for all $i = 1, ..., n-1$.

To find A, we first construct two matrices X_1 and X_2 by shifting the data points in X by one time step:

$$X_1 = \begin{bmatrix} x_1 & x_2 & \cdots & x_{n-1} \end{bmatrix}$$

$$X_2 = \begin{bmatrix} x_2 & x_3 & \cdots & x_n \end{bmatrix}$$

Next, we perform the Singular Value Decomposition (SVD) on X_1:

$$X_1 = U\Sigma V^*$$

where U and V are unitary matrices, and Σ is a diagonal matrix containing the singular values of X_1. We can then truncate the SVD by keeping only the first r singular values and corresponding columns of U and V. The reduced matrices U_r, Σ_r, and V_r can be used to approximate the matrix A:

$$A \approx U_r^* X_2 V_r \Sigma_r^{-1}$$

The eigenvalues and eigenvectors of A can then be computed to obtain the DMD modes and frequencies, which represent the dominant spatiotemporal patterns of the system's dynamics. The DMD modes can be used to reconstruct the system's evolution over time, and tasks such as background subtraction in video streams.

3 Results

The DMD removed the background variations from the video clip that was provided. It highlighted the foreground objects which are what we are interested in learning about. This is an integral step in detecting the track of the object of interest.

4 Conclusion

The DMD was successful in subtracting the background from the video clip. Overall then, the DMD algorithm presented here takes advantage of low dimensionality in the data in order to make a low-rank approximation of the linear mapping that best approximates the dynamics of the data collected from the video clip provided.

5 Acknowledgment

I received some helpful guidance from the TAs throughout this project. They provided useful insights on how to approach certain problems and were always available to answer any questions I had. I appreciate their support and am grateful for the assistance they provided.