Coding Project 5: Background Subtraction through Dynamic Mode Decomposition

Justin Hexem

3/17/2023

Abstract

In this project, we will separate the background and the foreground of a video. This was accomplished through the use of a technique called dynamic mode decomposition. This technique involves approximating a dynamical system with nothing but the time series data itself. This allows us to isolate the moving section of the image while discarding all of the stationary background parts of the image.

Introduction Dynamic mode decomposition is a method which takes advantage of low-dimensionality in complex systems in order to reduce a given system into a form that may be simpler and easier to work with. Most dimensionality reduction algorithms make use of an approach that is based on using the governing equations for that system to model the dynamical behavior of the system. Dynamic mode decomposition gives a different approach that allows us to model the dynamical behavior of a system without any knowledge of the governing equations of the system. It does this through the use of the data itself to understand the system.

We will make use of this method to separate the dynamic aspects of a video from the static background of the video by assuming that the frames of our video represent the time series for some dynamical system. This will allow us to use dynamic mode decomposition to track the movement in this video and isolate this movement from the rest of the video. This will be carried out in a MATLAB script that will process the input video and output the isolated foreground and background of our original videos.

Theoretical Background The procedure for dynamic mode decomposition begins by arranging our time series data into an nxm matrix X. This matrix is our original data with the spacial snapshots at each time point as column vectors. We will take X and form the matrices X_1^{m-1} and X_2^m such that X_1^{m-1} is the matrix X with the last column removes and X_2^m is the matrix X with the first column removed. This is done in order to approximate an operator called the Koopman operator, which is a linear, time-independent operator A that satisfies

$$x_{j+1} = Ax_j \tag{1}$$

where j represents time this data was collected and $j+1$ is the next time that the data is collected. We then take the singular value decomposition of $X_1^{m-1} = U\Sigma V^*$ and use this decomposition to form a new matrix

$$\tilde{S} = U^* X_2^m V \Sigma^{-1}. \tag{2}$$

Finding the eigenvectors and eigenvalues of $\tilde{S}$ gives a low-rank approximation that captures the dynamics of the Koopman operator A for one time step. From this, we can now obtain the dynamic mode decomposition modes by projecting the initial state of the system onto the dynamic mode decomposition modes with the following calculation

$$\psi_k = Uy_k. \tag{3}$$

We can then re-scale the eigenvalues of our $\tilde{S}$, denoted μ_k, to obtain $\omega_k = \ln(\mu_k)/\Delta t$. Now we may apply this approximation of the Koopman operator to our initial condition to step forward in time as many time periods as we would like. The formula for this is given by

$$x_{DMD}(t) = \sum_{k=1}^{K} b_k(0)\psi_k(x)e^{\omega t} \tag{4}$$

where $b_k(0)$ is the starting amplitude of each mode.

Results When we applied dynamic mode decomposition to the video, we were able to isolate the background of the video by taking the eigenvector corresponding to the smallest eigenvalue in magnitude and only using this vector to step forward in time. This means that this smallest eigenvector is the direction that the data changes the least, and when we examine the part of the video that remains mostly unchanged we see that this is the background of the video itself. Once we obtain this background, we can then just subtract it from our original data to obtain nothing but the foreground of the original video. This foreground is the part of the video that moves and changes the most, so it makes sense that the magnitude of the eigenvalues associated with these aspects of the video would be larger in magnitude.

Conclusion We were able to make use of dynamic mode decomposition to isolate and remove the background of a video by examining what parts of the video changed the least from frame to frame. This was done by applying this method in a MATLAB script and having as an output the background and foreground of the original video. By viewing the final separated videos of the background and foreground, we can see that dynamic mode decomposition was able to completely isolate the moving object from the static background with a high degree of accuracy.

Acknowledgment I greatly appreciated all the help that Professor Rahman provided while I was working through this project.