PCA on the Motion Data of a Cat Pounce

1. Introduction

This dataset was acquired by taking a 1920x1080 pixel 30 FPS video of two cats, Piccolo and Gohan. The video shows Piccolo standing still and Gohan running and pouncing at an object out of frame. We will use SVD and PCA to see how the data of Piccolo standing still and the motion of Gohan pouncing will appear on the score plot and loading plot reconstructions of the data.

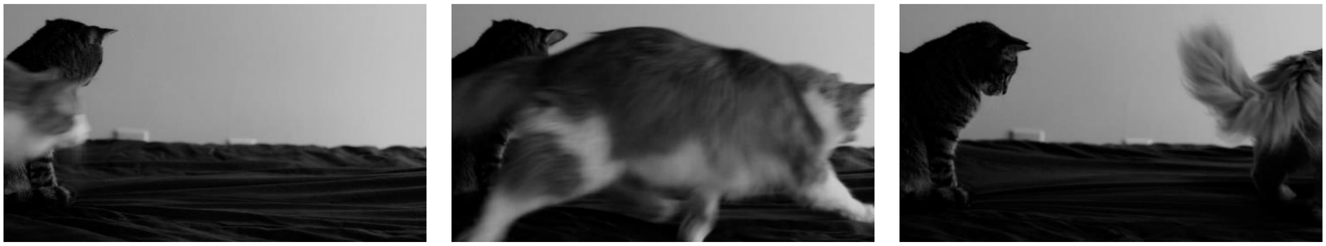

Figure 1. Frames 1, 11, and 24 of the Sequence

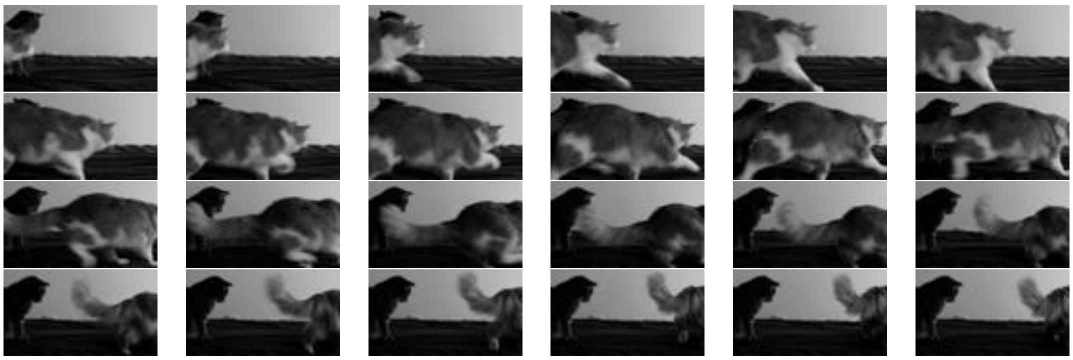

Figure 2. Frames 1 through 24 of the Sequence resized to 80x45 pixels (Left to Right, Top to Bottom)

2. Method

Singular Value Decomposition (SVD) will be performed on the dataset following Principal Component Analysis (PCA) to better understand the variance of the video frames.

I. Singular Value Decomposition

SVD is a process that decomposes a matrix into three component matrices: $U, \Sigma, and\ V$. The original matrix, A, can be rebuilt after being decomposed by taking the products of the component matrices, with the transpose of matrix V, in the form $A = U\Sigma V^T$. Each component of the decomposition can help to gain insights into the original matrix A through PCA.

II. Principal Component Analysis

The columns of matrix U are the left singular vectors, also known as the Principal Components (PCs) of A. These PCs are used to describe the amount of variance in the data contained in A. The columns of U can also be used to generate loading plots of the PCs, where the number of PCs are equivalent to the number of singular values in Σ. The loading plots of the individual PCs will show which pixels have the greatest influence on the corresponding pixels of frames in A. Σ contains the singular values, which are the square root of the eigenvalues, and are stored diagonally across the matrix. V contains the right singular vectors. The product of ΣV^T generates the score matrix, which can be used to create a score plot of two PCs – usually the first and second PCs. The score plot shows the interaction between two PCs, or two amounts of variance, and the original data.

3. Data

The image data contains 24 observations as columns, 3600 features as rows, and labels were generated for each image. The observations of the data are the 24 columns, which represent each image frame of the video. The features of the data are the 3600 rows, each of which are an individual, 1-dimensional representation of the 24 80x45 pixel frames of the video. In addition to the preprocessing of reducing the frame sizes down to 80x45 pixels, the data was centered.

Preprocessing

I. Grayscale, Resizing, and Flattening

The video was converted to grayscale to remove the RGB channels and allow for easier creation of an image matrix containing only the black levels. To decrease the computational time of the code, the original video was trimmed down from 70 to 24 frames (0.8 seconds) of the original video and resized by a factor of (coincidentally) 24 to 80x45 pixels. The 24 images were then flattened by concatenating the rows of pixel data for each image to a single vector of length 3600. The flattened images were then added sequentially to an array of 3600 columns and 24 rows.

II. Centering

Mean-centering was done along the rows of the data by taking the mean of each row (image) and subtracting the mean from each value in the row. Centering the data creates better results for the SVD process causing the score matrix of the principal components to be centered about the origin.

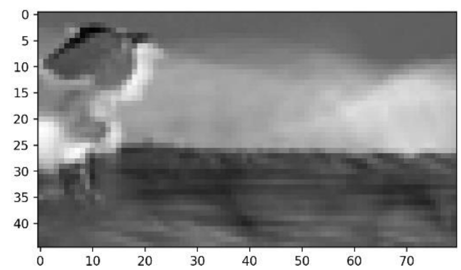

Figure 3. Centered Data for Frame 1

4. Results

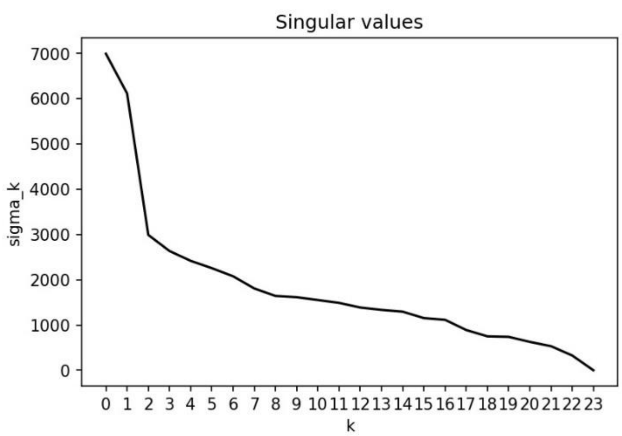

Figure 4. Singular Values of the image data SVD

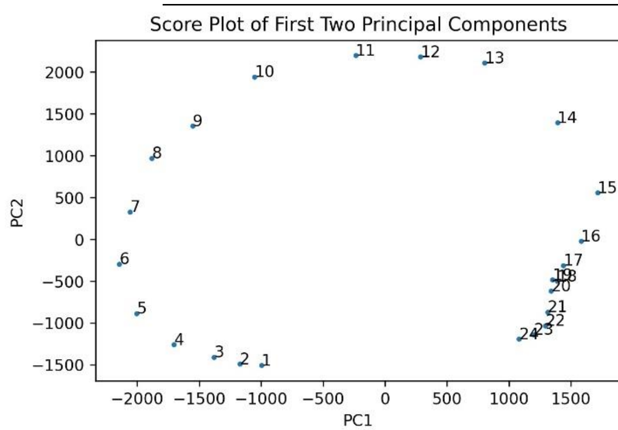

Figure 5. Score Plot of The First 2 Principal Components

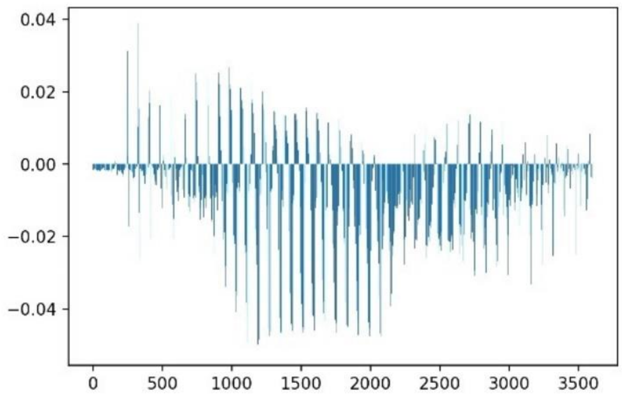

Figure 6. Loading Plot for the First Principal Component as a Bar Graph

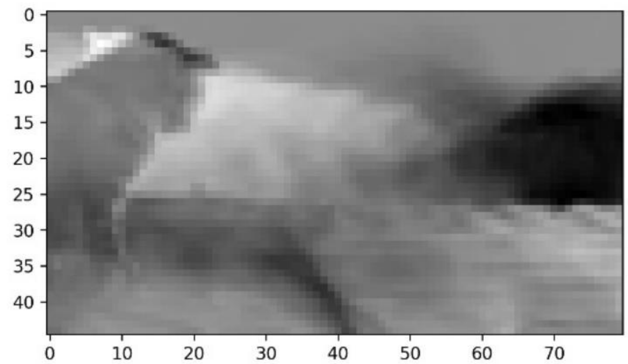

Figure 7. Loading Plot for the First Principal Component as an Image

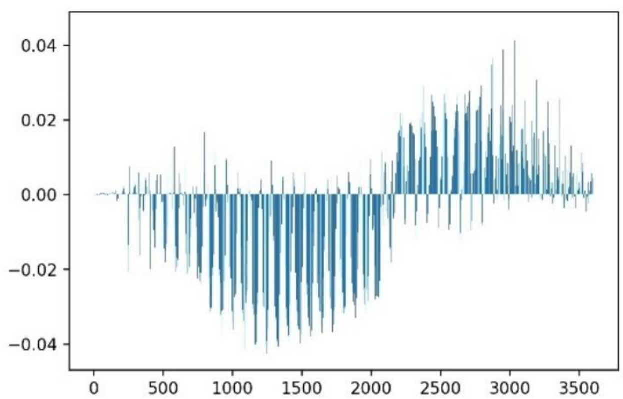

Figure 8. Loading Plot of the Second Principal Component as a Bar Graph

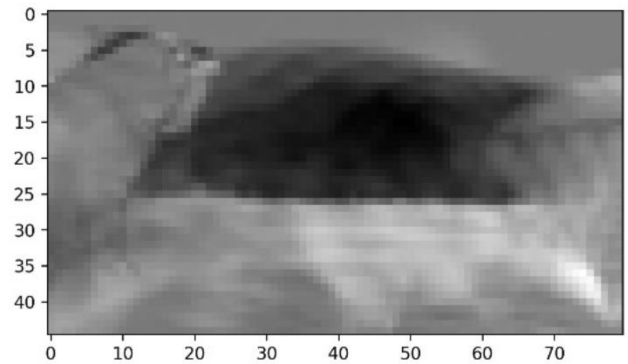

Figure 9. Loading Plot for the Second Principal Component as an Image

5. Discussion

Principal Components:	First PC	First Two PCs	First Three PCs	First Four PCs	First Five PCs
Variation Explained:	34.622%	61.091%	67.421%	72.336%	76.474%

Figure 10. Variation Explained up to First Five PCs

Figure 10 shows that the first PC only explains 34.622% of the variance, while the first two PCs explain 61.091% of the variance. It takes a combination of the first five PCs to explain 76.474% of the variance. The variance explains the motion, or the change in pixel data, between the 24 frames. The variance mostly describes the motion of Gohan, as Piccolo stays relatively still throughout, with his outline visible on the left of the loading plots shown in Figures 8 and 9 for PC1 and PC2, respectively.

The score plot of PC1 and PC2 in Figure 5 produces an interesting graph due to the nature of the motion data. The score plot is "circular", or appears to create a clockwise cycle, with the data of Frame 1 beginning the cycle, and the data of Frame 24 almost completing the cycle. Frame 24 does not fully complete the cycle because it is only slightly similar to, but not exactly the same as the image in Frame 1: Gohan's head pokes out from the left of Frame 1 and his tail is poking out from the right of Frame 24. The score plot data for Frames 15 through 24 are also more clustered together while the other points are relatively equally spaced. This is due to Gohan running through the frames at a relatively constant speed but starting to slow down around Frame 15 and stopping towards the end of the sequence while still being in the image.

The loading plots and their image representations in Figures 6 through 9 show the amount of variation between each pixel described by each principal component. The dark and light blurs show where more motion, or variation, occurred. This can be seen by the dark and light spots around Piccolo's head in Figures 7 and 9, as well as the dark and light spots where pixels of Gohan's body intersected with itself as the frames progressed.

6. Evaluation

PCA was an appropriate tool to use in analyzing the motion data, as the score plot displayed a representation of Gohan's motion throughout the sequence. In addition to this finding, it is possible that a rough estimate of Gohan's position, velocity, and acceleration over the sequence could be obtained by finding the distance between frame's point in the score plot.

This analysis could be improved by identifying the shapes of each cat, and taking the combined 3600 most important pixels, where importance would be defined by the variance in the shape outlines, of each cat. The frames could then be parsed to identify the changes in those outline pixels throughout the sequence. The motion sequence of those 3600 combined outline pixels could then have PCA performed on it to get a better understanding of the variance in Piccolo's minor head motion and Gohan's pouncing motion.