Motion Detection Using Dynamic Mode Decomposition

Marco Mignacca[1], Simone Brugiapaglia[2], Jason Bramburger[2]

1: Department of Mathematics and Statistics, McGill University, 2: Department of Mathematics and Statistics, Concordia University

Introduction

- **Dynamic Mode Decomposition (DMD)** is a simple, interpretable data-driven technique to disentangle space-time dynamics
- Provides a linear approximation to a dynamical system
- Input: snapshots arranged in two **data matrices**

$$X = [X_1 \quad X_2 \quad \cdots \quad X_n],\ Y = [Y_1 \quad Y_2 \quad \cdots \quad Y_n],\ Y_i = X_{i+1}$$

- **DMD matrix** $A = YX^{\dagger}$ solves:

$$\underset{A \in \mathbb{C}^{M \times M}}{\operatorname{argmin}} \|Y - AX\|_F^2$$

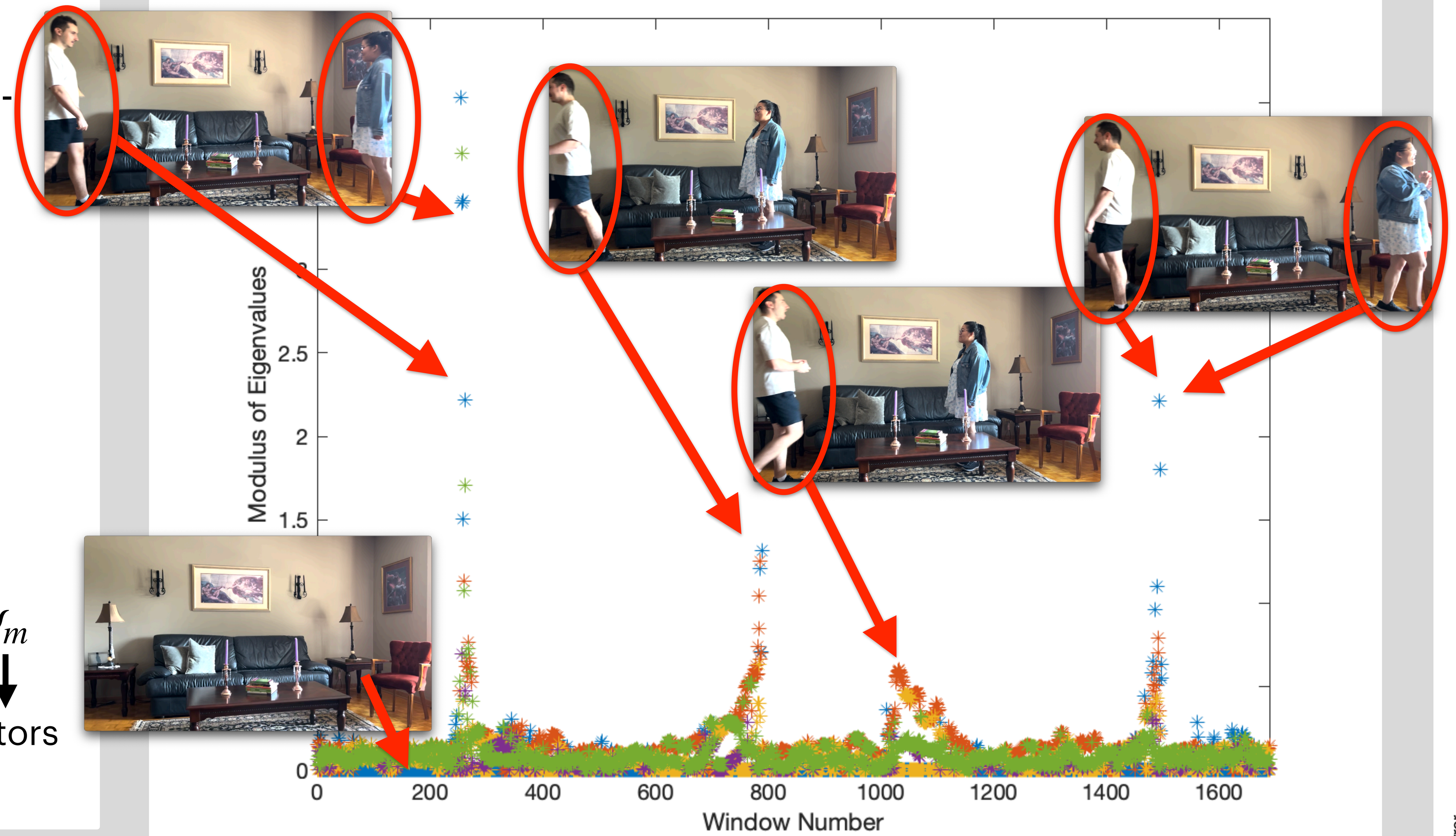
If A is diagonalizable, we can decompose: $X_k = \sum_{m=1}^{M} c_m \lambda_m^k \psi_m$

Eigenvalues

Eigenvectors

Objective: leverage this to implement an automatic motion detection system for video

Motion Detection Scheme



Eigenvalue average in each window

$$\mu^{(k)} = \frac{1}{r}\sum_{i=1}^{r} |\omega_i^{(k)}|$$

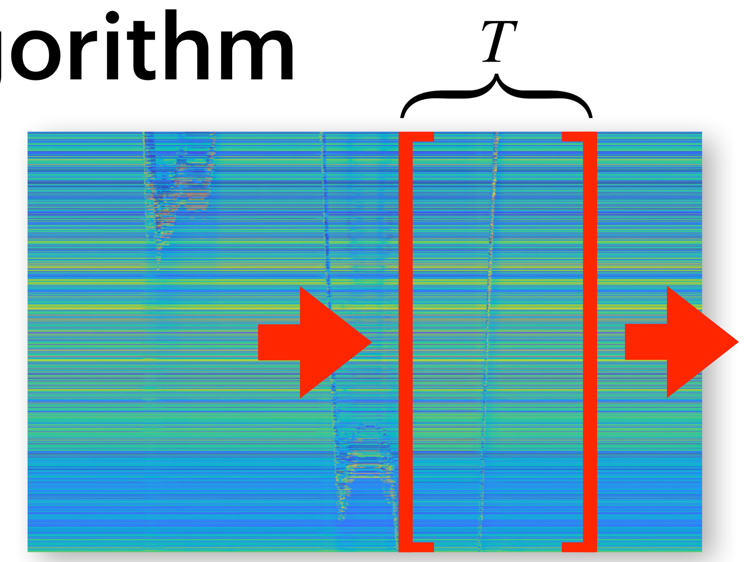
Compute relative change

$$\Delta = \frac{|\mu^{(k+1)} - \mu^{(k)}|}{|\mu^{(k)}|}$$

$\Delta > \Delta^*$ → Motion Detected!

Optimizing Δ^* requires an error measure: $E = FP + c \cdot FN$

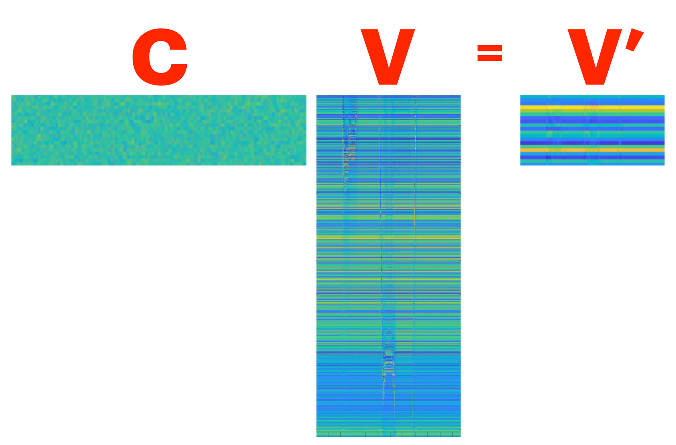
DMD Algorithm



Input: video matrix $V \in \mathbb{R}^{M \times N}$

1. Compressed DMD

High dimensional video data V is compressed using a random matrix $C : V' = CV$



2. Sliding Window DMD

Streaming data is separated into windows of size T to apply DMD

3. Rank 'r' Approximation

Take the SVD of compressed data matrix X' to find best rank r approximation: $X' \approx U_r \Sigma_r V_r^*$

4. Rank Reduced DMD Matrix

Calculate the projection of the DMD matrix A onto U_r:

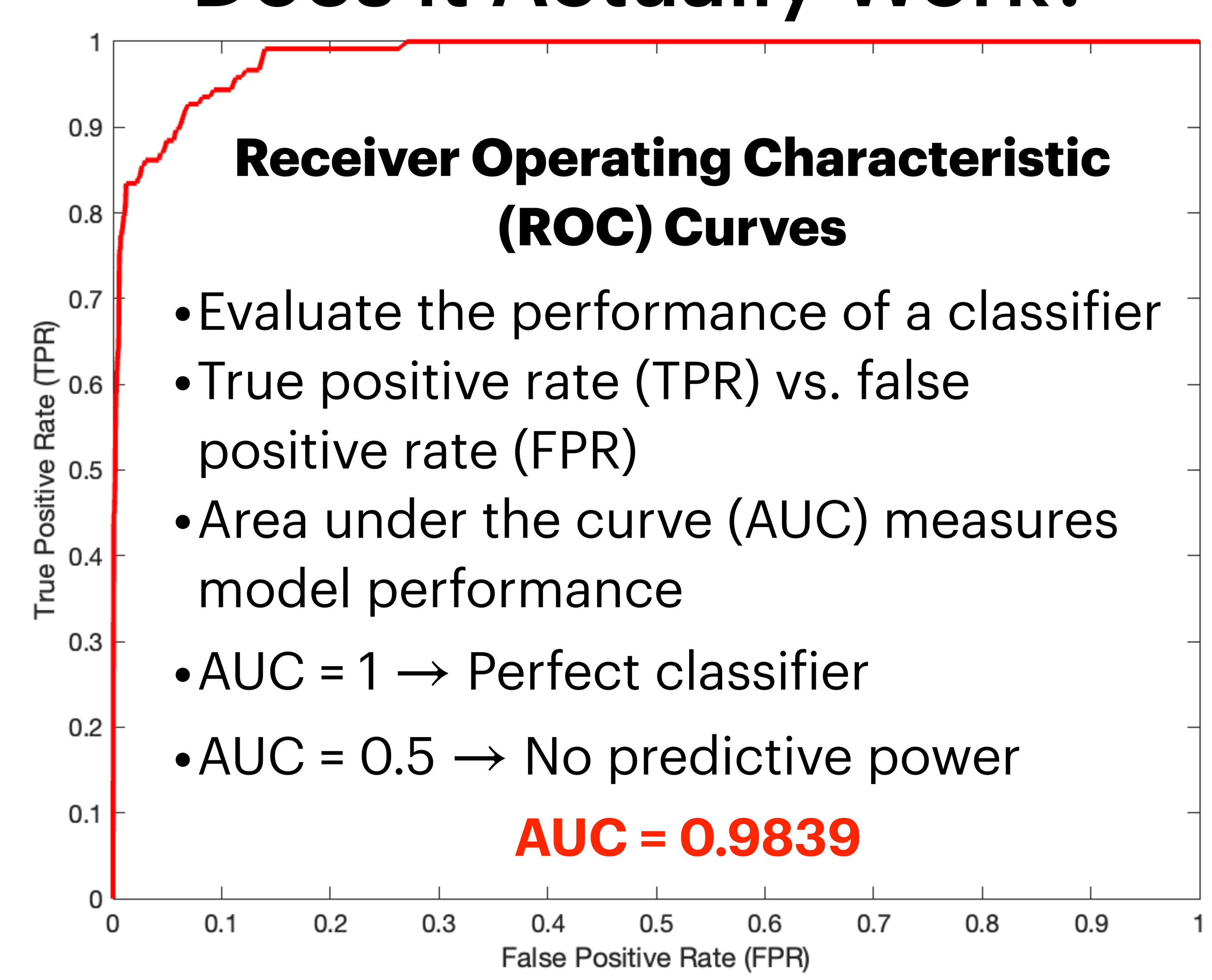
$$\tilde{A} = U_r^* A U_r = U_r^T Y_2 V_r \Sigma_r^{-1}$$

5. Extract Eigenvalues

$$\tilde{A}v = \lambda v \implies \omega = \log \lambda / \Delta t$$

Output: matrix of eigenvalues

Does it Actually Work?



Receiver Operating Characteristic (ROC) Curves

- Evaluate the performance of a classifier
- True positive rate (TPR) vs. false positive rate (FPR)
- Area under the curve (AUC) measures model performance
- AUC = 1 → Perfect classifier
- AUC = 0.5 → No predictive power

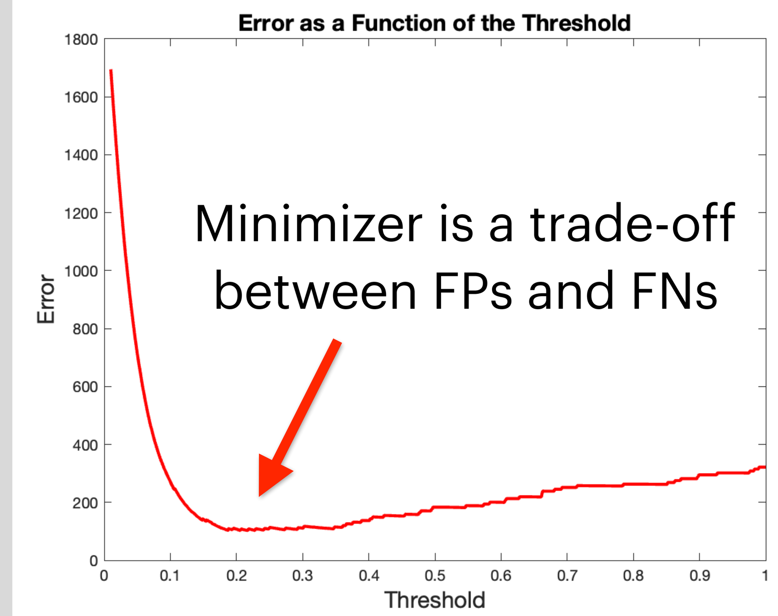
AUC = 0.9839

Event Detection Training

Pseudo k-fold Cross Validation

1	V	T	T	T
2	T	V	T	T
3	T	T	V	T
4	T	T	T	V

1. Shuffle the given dataset randomly
2. Split into k groups (**folds**) of equal size
3. For k iterations
   1. Separate into **training set (T)** and **validation set (V)**
   2. Train the model on T, tracking error for each Δ^*
   3. Threshold with smallest average error is optimal
   4. Test the threshold value on V and compute error
5. Threshold with lowest overall error is optimal



Results

		Iter 1	Iter 2	Iter 3	Iter 4
Easy	Threshold	0.450	0.348	0.410	0.348
	Avg. err.	17.00	31.33	33.67	68.33
Hard	Threshold	0.348	0.348	0.360	0.450
	Avg. err.	91.00	29.00	28.80	127.00

Conclusion

We have proposed a simple, effective, and interpretable method to detect movement in video

- Method works best when subjects are clear and isolated
- Future work: comparison with standard motion detection methods and training with a larger video database

References

1. Daniel Dylewsky, Molei Tao, and J. Nathan Kutz. "Dynamic mode decomposition for multiscale nonlinear physics". In: Phys. Rev. E 99 (6 June 2019), p. 063311.
2. Erichson, N.B., Brunton, S.L. & Kutz, J.N. Compressed dynamic mode decomposition for background modeling. J Real-Time Image Proc 16, 1479–1492 (2019).
3. Gareth James et al. An Introduction to Statistical Learning with Application in R. 2nd ed. Springer Texts in Statistics. Springer New York, NY, July 2021, pp. XV–607.
4. J. Nathan Kutz et al. Dynamic Mode Decomposition. Society for Industrial and Applied Mathematics, 2016.

This research was funded by NSERC (Undergraduate Student Research Award) and Concordia University

marco.mignacca@mail.mcgill.ca