Due Friday March 17th at 11:59pm

Background Subtraction in Video Streams

Use dynamic mode decomposition on the video clip `phi4.mov` from my paper "A mechanical analog of the two-bounce resonance of solitary waves: modeling and experiment". *Chaos* **25**, 043109 (2015); DOI: 10.1063/1.4917047. In this paper we used a different background subtraction algorithm because we didn't know about Dynamic Mode Decomposition.

The DMD spectrum of frequencies can be used to subtract background modes. Specifically, assume that ω_p, where $p = \{1, 2, \ldots, l\}$, satisfies $\|\omega_p\| \approx 0$, and that $\|\omega_j\| \forall j \neq p$ is bounded away from zero. Thus,

$$\mathbf{X}_{\text{DMD}} = \underbrace{b_p \varphi_p e^{\omega_p \mathbf{t}}}_{\text{Background Video}} + \underbrace{\sum_{j \neq p} b_j \varphi_j e^{\omega_j \mathbf{t}}}_{\text{Foreground Video}} \tag{1}$$

Assuming that $\mathbf{X} \in \mathbb{R}^{n\times m}$, then a proper DMD reconstruction should also produce $\mathbf{X}_{\text{DMD}} \in \mathbb{R}^{n\times m}$. However, each term of the DMD reconstruction is complex: $b_j \varphi_j \exp(\omega_j \mathbf{t}) \in \mathbb{C}^{n\times m} \forall j$, though they sum to a real-valued matrix. This poses a problem when separating the DMD terms into approximate low-rank and sparse reconstructions because real-valued outputs are desired and knowing how to handle the complex elements can make a significant difference in the accuracy of the results. Consider calculating the DMD's approximate low-rank reconstruction according to

$$\mathbf{X}_{\text{DMD}}^{\text{Low-Rank}} = b_p \varphi_p e^{\omega_p \mathbf{t}}.$$

Since it should be true that

$$\mathbf{X} = \mathbf{X}_{\text{DMD}}^{\text{Low-Rank}} + \mathbf{X}_{\text{DMD}}^{\text{Sparse}},$$

then the DMD's approximate sparse reconstruction,

$$\mathbf{X}_{\text{DMD}}^{\text{Sparse}} = \sum_{j \neq p} b_j \varphi_j e^{\omega_j \mathbf{t}},$$

can be calculated with real-valued elements only as follows:

$$\mathbf{X}_{\text{DMD}}^{\text{Sparse}} = \mathbf{X} - |X_{\text{DMD}}^{\text{Low-Rank}}|,$$

where $|\cdot|$ yields the modulus of each element within the matrix. However, this may result in $\mathbf{X}_{\text{DMD}}^{\text{Sparse}}$ having negative values in some of its elements, which would not make sense in terms of having negative pixel intensities. These residual negative values can be put into a $n \times m$ matrix $\mathbf{R}$ and then be added back into $\mathbf{X}_{\text{DMD}}^{\text{Low-Rank}}$ as follows:

$$\mathbf{X}_{\text{DMD}}^{\text{Low-Rank}} \leftarrow \mathbf{R} + |\mathbf{X}_{\text{DMD}}^{\text{Low-Rank}}$$
$$\mathbf{X}_{\text{DMD}}^{\text{Sparse}} \leftarrow \mathbf{X}_{\text{DMD}}^{\text{Sparse}} - \mathbf{R}$$

This way the magnitudes of the complex values from the DMD reconstruction are accounted for, while maintaining the important constraints that

$$\mathbf{X} = \mathbf{X}_{\text{DMD}}^{\text{Low-Rank}} + \mathbf{X}_{\text{DMD}}^{\text{Sparse}}.$$

so that none of the pixel intensities are below zero, and ensuring that the approximate low-rank and sparse DMD reconstructions are real-valued. This method seems to work well empirically.