Analysis of Periodicity in Video Sequences Through Dynamic Linear Modeling

A. Jonathan McLeod[1,2(✉)], Dante P.I. Capaldi[1], John S.H. Baxter[1,2], Grace Parraga[1], Xiongbiao Luo[3], and Terry M. Peters[1,2]

[1] Robarts Research Institute, Western University, London, Canada
jmcleod@robarts.ca
[2] Biomedical Engineering Graduate Program, Western University, London, Canada
[3] Department of Computer Science, Xiamen University, Xiamen, China

Abstract. Periodicity is an important characteristic in many types of video sequences, particularly in medical applications where the cardiac and respiratory cycles are of special significance. Simple spectral analysis or band-pass filtering is often insufficient to extract the periodic signal. Here, we propose modeling the periodic and background components using nested dynamic linear models. These models can approximate the periodic and background time series in a wide range of video sequences. A likelihood ratio test can be used to find regions of the video exhibiting periodicity. Our experiments suggested this technique is suitable for a variety of applications using different imaging modalities, including ultrasound, MRI and natural video.

Keywords: Periodicity detection · Dynamic linear models · Remote photoplethysmogram · Perfusion mapping · Fourier-decomposition MRI

1 Introduction

Analyzing quasiperiodic variations in a video sequence is frequently performed in medical imaging with the goal of extracting information related to the cardiac or respiratory cycles. Natural video of a person's face or hand can be used for non-contact monitoring of vital signs including heart rate, respiratory rate, and pulse transit time [10,11] or to generate maps showing the spatial distribution of tissue perfusion [7,12]. Fourier-decomposition MRI is an emerging technique for imaging lung perfusion and ventilation that relies on Fourier analysis of a sequence of non-contrast enhanced MR images [3]. Another area where periodicity is useful in medical imaging is in the detection of critical pulsating structures in medical interventions [1,2,8,9]. Most of these techniques require extensive filtering and preprocessing that have been fine-tuned for their respective problems. This is necessary to produce a sufficiently clean periodic signal that can be extracted through Fourier analysis or band-pass filtering.

In this paper we propose the use of dynamic linear modeling for analyzing periodicity in video sequences. We show how a cyclic + random walk model can

be used to estimate the frequency and amplitude of quasi-periodic components. We also propose a log-likelihood ratio statistic for determining the presence of periodicity. In addition, we derive the power spectral density function for this model and show that it closely resembles that of the observed spectrum commonly found in video sequences. This approach is applied to natural video, ultrasound and MRI.

2 Methods

2.1 Dynamic Linear Models

Dynamic Linear Models (DLM) are linear state-space time series models [6] of the form,

$$y_t = Za_t + \epsilon_t, \qquad \epsilon_t \sim N(0, \Sigma_\epsilon), \qquad a_t = Ta_{t-1} + \eta_t, \qquad \eta_t \sim N(0, \Sigma_\eta), \quad (1)$$

where y_t and a_t are the observation and state vectors at time t. The observation and state transition matrices are Z and T with independent normally distributed noise covariances Σ_ϵ and Σ_η. Often these are block diagonal matrices formed from several simpler models whose parameters have an intuitive or physical meaning. Maximum likelihood estimates (MLE) for unknown parameters in Z, T, Σ_ϵ and Σ_η can be obtained by numerically optimizing the log-likelihood function,

$$\log L(\theta|Y_t) = \sum \log p(y_t|Y_t, \theta), \tag{2}$$

where Y_t denotes the vector of observations up to time t, θ are the unknown model parameters and $p(\cdot)$ is the probability density function. Kalman filtering is used to evaluate $\log L(\theta|Y_t)$.

We propose a nested DLM to model quasi-periodicity in video sequences. This model consists of a stationary cyclic component, random walk component and additive measurement noise as specified in Eq. 3,

$$Z = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}, \quad T = \begin{bmatrix} \rho\cos\omega_0 & \rho\sin\omega_0 & 0 \\ -\rho\sin\omega_0 & \rho\cos\omega_0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Sigma_\epsilon = \sigma_n^2, \quad \Sigma_\eta = \begin{bmatrix} \sigma_c^2 & 0 & 0 \\ 0 & \sigma_c^2 & 0 \\ 0 & 0 & \sigma_l^2 \end{bmatrix}. \tag{3}$$

The parameters ω_0 and ρ represent the frequency and bandwidth of the cyclic component. The variances σ_c^2, σ_l^2 and σ_n^2 specify the strength of cyclic, random walk and additive noise components respectively. In the state vector, $a_t = [a_{t,1}, a_{t,2}, a_{t,3}]^T$, the first two states, $a_{t,1}$, and $a_{t,2}$ are analogous to the real and imaginary components of a complex oscillator, while the third state, $a_{t,3}$ follows a random walk to account for signal drift and other low frequency variations. The model observation matrix, Z, adds $a_{t,1}$ and $a_{t,3}$ to obtain a cyclic model with a local level that follows a random walk.

The power spectral density of the model can be derived by breaking apart the cyclic and random walk components. Transfer functions from the cyclic component state noise $\eta_{t,1}$ and $\eta_{t,2}$ to the output y_t can be obtained directly from

the state-space model:

$$G_1(z) = \frac{z - \rho\cos(\omega_0)}{z^2 - 2\rho\cos(\omega)z + \rho^2}, G_2(z) = \frac{-\rho\sin(\omega_0)}{z^2 - 2\rho\cos(\omega)z + \rho^2}, \qquad (4)$$

where $G_1(z)$, $G_2(z)$ are the respective transfer functions and z is the forward shift operator. The power spectral density can be obtained by evaluating the transfer functions at $z = e^{j\omega}$ yielding,

$$S_{yy}^{cyclic}(\omega) = \frac{\sigma_c^2(1 + \rho^2 - 2\rho\cos\omega_0\cos\omega)}{[1 + \rho^2 - 2\rho\cos(\omega - \omega_0)][1 + \rho^2 - 2\rho\cos(\omega + \omega_0)]}. \qquad (5)$$

The random walk is non-stationary with a well known frequency drop-off approximately proportional to ω^{-2}. An exact expression for its power spectral density, $S_{yy}^{rw}(\omega)$, with a finite length time series can be obtained through Fourier analysis where the Fourier transform of the random walk state, $a_{t,3}$ is expressed in terms of the incremental state noise, $\eta_{t,3}$:

$$F(a_{t,3}) = \sum_{i=0}^{N} F(\eta_{i,3}H(t - i)), \qquad (6)$$

where F is the discrete Fourier transform operator and $H(t)$ is the Heaviside function. Since $\eta_{t,3}$ are independent random variables, the power spectral density can be obtained from the expected value of Eq. 6 as,

$$S_{yy}^{rw}(\omega) = \mathbf{E}|F(a_{t,3})|^2 = \sum_{i=1}^{N} \mathbf{E}|F(\eta_{t,3}H(t - i))|^2 = \frac{\sigma_l^2}{1 - \cos\omega}. \qquad (7)$$

The total power spectral density for Eq. 3 is obtained from Eqs. 5 and 7 plus the white additive measurement noise,

$$S_{yy}(\omega) = \frac{\sigma_c^2(1 + \rho^2 - 2\rho\cos\omega_0\cos\omega)}{[1 + \rho^2 - 2\rho\cos(\omega - \omega_0)][1 + \rho^2 - 2\rho\cos(\omega + \omega_0)]} + \frac{\sigma_l^2}{1 - \cos\omega} + \sigma_n^2. \qquad (8)$$

As can be seen in Fig. 1, the power spectral density of the model slopes downward at -20 dB/decade with a small peak near ω_0 corresponding to the cyclic component. This closely resembles the empirical spectrum observed in Figs. 2 and 4. When $\sigma_c^2 \to 0$, the cyclic component vanishes and we are left with just the random walk and additive noise also known in structural time series as the local level model. Thus, to test for the presence of a cyclic component, we can test $\mathcal{H}_0 : \sigma_c^2 = 0$ against simple negation. Let $\hat{\theta}_0$ denote the MLE in the restricted model, $\sigma_c^2 = 0$. Then, if $\mathcal{H}_0$ is correct, the likelihood-ratio test statistic,

$$D = -2[\log L(\hat{\theta}|Y_t) - \log L(\hat{\theta}_0|Y_t)], \qquad (9)$$

measures the relative plausibility of $\mathcal{H}_0$. We use D to test for periodically varying regions in the video.
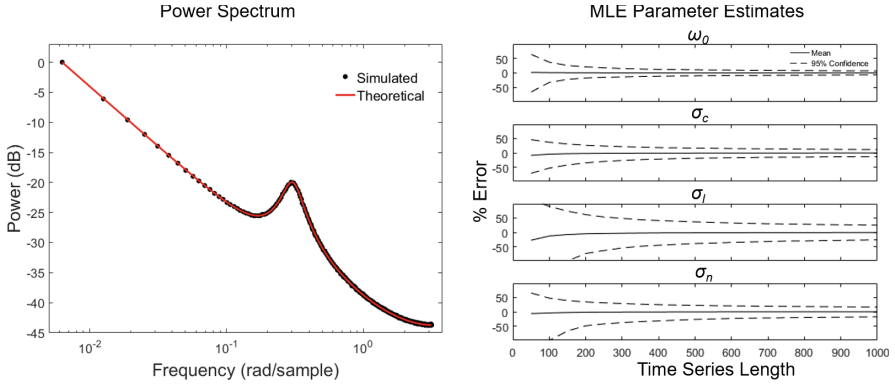
Fig. 1. The analytic spectrum of the DLM given in Eq. (8) and convergence of the MLE of model parameters are verified through numerical simulations. These graphs are shown for $\omega_0 = 0.3$, $\rho = 0.95$, $\sigma_c / \sqrt{(1 - \rho^2)} = 5\sigma_l = 5\sigma_n$

2.2 Experiments

To demonstrate the effectiveness of this model, we considered three very different datasets consisting of MRI, ultrasound and natural video. The first dataset consisted of natural video of a human hand. We demonstrate how DLM can be used to estimate the frequency (heart rate) and amplitude (perfusion map) from the video. Next we apply DLM to an ultrasound video of the lumbar spine where pixels in the dura exhibit subtle pulsation. Here we use the likelihood ratio statistic, D, to test for the presence of periodicity. Finally, we consider a free-breathing lung MRI sequence where local ventilation images can also be generated from DLM.

3 Results

3.1 Natural Video

Photoplethysmogram (PPG) is an optical measurement of cardiac activity. Typically, infrared light is used in pulse-oximeters due to better tissue penetration depth. However, the ubiquity of digital cameras has led to increasing interest in monitoring vital signs using ambient visible light. These systems have the advantage of monitoring vital signs remotely without requiring carefully controlled lighting or any direct contact with the patient. Methods have been proposed to extract heart rate, respiratory rate, and pulse transit time from videos of a person's face or hand [10,11]. Beyond simply measuring vital signs, it is also possible to generate spatial maps showing variations in the magnitude of the PPG signal [7,12]. PPG imaging has the potential to show tissue perfusion relevant to many clinical problems such as evaluating skin-flaps and burn injuries. These techniques typically require extensive preprocessing to detrend the PPG

signal and remove the effects of motion or variations in ambient light. For this experiment 8 videos were acquired showing the hands of seven subjects. Videos 1&2 were aquired of the same subject to demonstrate perfusion mapping. The hand was gently scratched between these aquisitions to stimulate blood flow. This test has been previously used for perfusion mapping techniques based on PPG imaging [7] and laser Doppler imaging [4].

First, we demonstrate that heart rate can be estimated from all 8 videos. A PPG signal was extracted by averaging the green channel intensity over a 400×400 block as shown in Fig. 2. The proposed DLM was fit to the PPG signal and the estimate of ω_0 compared with the Fourier spectral peak and the readings from the pulse oximeter (Table 1). The absolute error between DLM frequency estimation and the nearest pulse oximeter reading was 2.3 ± 1.0 bpm (mean $\pm$ standard error). This was significantly lower ($p < 0.05$) than that estimates through simple Fourier analysis or quadratic peak interpolation (QPI) of the spectral peak, which were 5.4 ± 1.3 and 3.8 ± 1.4 respectively.
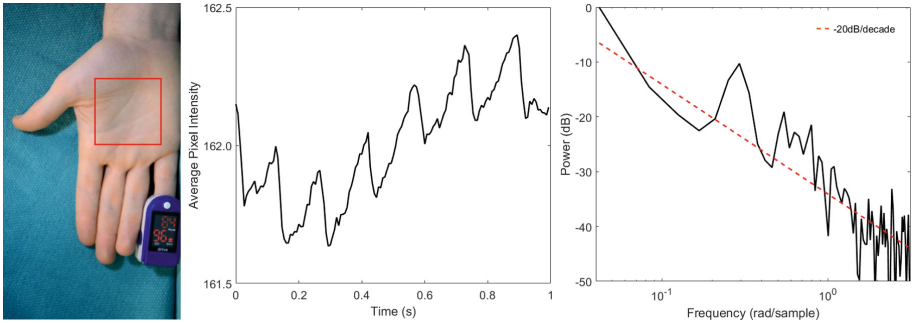

Fig. 2. Video clips were recorded of the subjects' hands and the PPG signal was calculated from the average intensity over the 400×400 box. The -20 dB/decade slope observed in the spectrum is consistent with the proposed DLM.

Table 1. Heart rate estimated from the 8 video clips (bpm)

	1	2	3	4	5	6	7	8	Abs Err
DLM	84.1	79.6	86.6	89.1	73.6	78.5	55.1	63.3	2.3 ± 1.0
Fourier	84.0	72.0	84.0	84.0	72.0	84.0	60.0	60.0	5.4 ± 1.3
QPI	82.7	76.4	83.7	86.9	70.7	82.8	55.7	63.1	3.8 ± 1.4
Pulse Ox.	84	82	94–98	90–91	79	80–81	56	63–65	

Next, to generate perfusion maps, the DLM is fitted on a per pixel level for videos 1&2. Here, the amplitude of the cyclic component corresponds to perfusion. The original 1080×1920 videos were reduced to 135×240 by applying a Gaussian blur with $\sigma = 8$ and a down-sampling factor of 8. The DLM was

fit with the frequency fixed to the MLE estimate given in Table 1. The quantity $\sigma_c/(\sqrt{1-\rho^2})$, corresponding to the amplitude of the cyclic component, is shown in Fig. 3. For comparison, the Fourier estimates of amplitude are also shown.
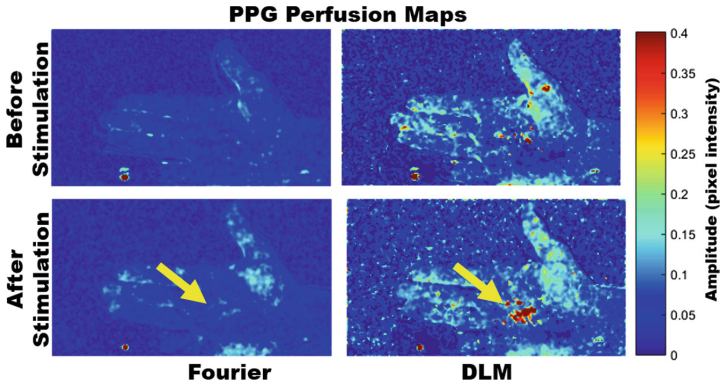
Fig. 3. PPG imaging using Fourier analysis and DLM. The yellow arrow indicates where the hand was scratched.

Without any preprocessing or tuning for this specific problem, a relatively simple DLM was able to identify very subtle changes in tissue perfusion that occurred after gently scratching the hand. Furthermore, this was accomplished with very short video clips only 5 s in length unlike previous methods that require much longer video clips and extensive preprocessing and detrending.

3.2 Ultrasound

Dural pulsation is a valuable cue in ultrasound guided epidural injections. Previously McLeod et al. [8] proposed an extended Kalman filtering (EKF) method that estimated frequency and amplitude of the pulsating dura on a per-pixel basis in lumbar spine ultrasound. Here, our main objective is to identify which pixels in the image exhibit periodicity. The likelihood-ratio statistic in Eq. 9 is ideal for this purpose. We fit the proposed DLM on a per-pixel basis to a video of the lumbar ultrasound and compared it against those obtained from the EKF method in McLeod et al. [8]. The results are nearly identical despite the EKF having being developed for this application only, and requiring extensive smoothing and tight thresholds on the frequency and amplitude (Fig. 4).

3.3 MRI

Fourier-decomposition of free-breathing proton MRI (FDMRI) has recently emerged as a non-contrast enhanced MRI technique to generate regional pulmonary ventilation maps on any clinically available MRI system [3,5]. This technique exploits fast pulmonary MRI acquisition and non-rigid image registration
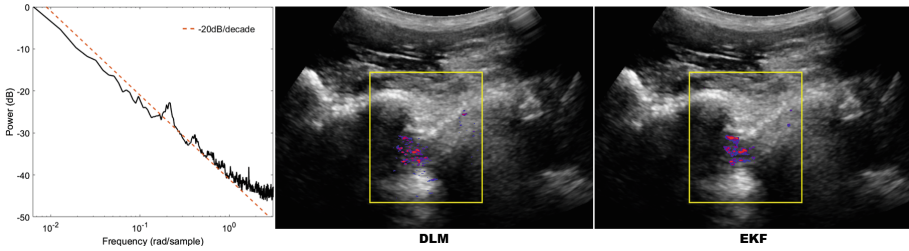
Fig. 4. The average power spectrum of the lumbar ultrasound ROI exhibits an approximately -20 dB/decade slope with a small peak at the cardiac frequency, closely resembling the theoretical DLM spectrum in Fig. 1. The likelihood ratio estimated from the DLM immediately reveals the location of the dural pulsation. The results are very similar to the hand-crafted EKF method [8].

to acquire a time series of registered proton MR images. Since the proton density within lung tissue varies with the respiratory cycle as the aveoli expand and contract, periodicity in the signal intensity of the registered images provides a measure of tissue ventilation. In FDMRI, the amplitude of the Fourier component at the respiratory frequency is used as a measure of lung ventilation. We acquired a dynamic free-breathing MRI of a non-small cell lung cancer patient over 125 s at a rate of 4 frames per second where the left lung was obstructed and poorly ventilated. A hyperpolarized ^{129}Xe MR static ventilation image was acquired as a benchmark and shows a lack of ventilation in the left lung. The DLM was fit to the MRI sequence (Fig. 5). The DLM amplitude map was visually similar to FDMRI, but with slightly better rejection of background tissue motion. The likelihood-ratio statistic provided a statistical test for the presence of ventilation. It showed a lack of ventilation in the left lung, only exhibiting motion artifacts around the lung boundary and was qualitatively closest to the ^{129}Xe MRI.
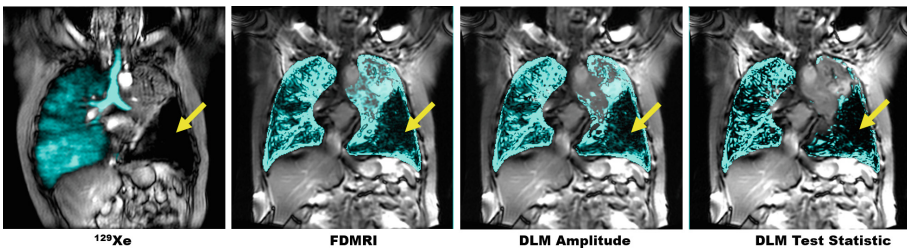
Fig. 5. The benchmark ^{129}Xe MR image is compared with FDMRI and the two DLM images (amplitude, and likelihood-ratio). The cyan overlays display the ventilation maps and the yellow arrows point to the lack of ventilation in the left lung. Of the three images generated from the free-breathing sequence, the DLM likelihood-ratio appears most similar to the ^{129}Xe MR image.

4   Discussion and Conclusion

We have shown how DLM provides a powerful framework for analyzing periodicity in various video sequences. The same DLM was applied directly to natural video, ultrasound and MRI without additional preprocessing or fine-tuning and it provided frequency and amplitude estimates as well as log-likelihood statistic testing for the presence of periodicity. The main strength of this model is that it provides a general and robust method for analyzing periodicity in video sequences that typically require individually handcrafted techniques. In addition, DLM can easily be extended to include multiple frequencies and harmonics as well as more advanced background noise models.

References

1. Amir-Khalili, A., Hamarneh, G., Abugharbieh, R.: Automatic vessel segmentation from pulsatile radial distension. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 403–410. Springer, Cham (2015). doi:10.1007/978-3-319-24574-4_48
2. Amir-Khalili, A., Hamarneh, G., Peyrat, J.M., Abinahed, J., Al-Alao, O., Al-Ansari, A., Abugharbieh, R.: Automatic segmentation of occluded vasculature via pulsatile motion analysis in endoscopic robot-assisted partial nephrectomy video. Med. Image Anal. **25**(1), 103–110 (2015)
3. Bauman, G., Puderbach, M., Deimling, M., Jellus, V., Chefd'hotel, C., Dinkel, J., Hintze, C., Kauczor, H.U., Schad, L.R.: Non-contrast-enhanced perfusion and ventilation assessment of the human lung by means of Fourier decomposition in proton MRI. Magn. Reson. Med. **62**(3), 656–664 (2009)
4. Briers, J.D.: Laser doppler, speckle and related techniques for blood perfusion mapping and imaging. Physiol. Measur. **22**(4), R35–R66 (2001)
5. Capaldi, D.P., Sheikh, K., Guo, F., Svenningsen, S., Etemad-Rezai, R., Coxson, H.O., Leipsic, J.A., McCormack, D.G., Parraga, G.: Free-breathing pulmonary ^{1}H and hyperpolarized ^{3}He MRI: comparison in COPD and bronchiectasis. Acad. Radiol. **22**(3), 320–329 (2015)
6. Durbin, J., Koopman, S.J.: Time Series Analysis by State Space Methods. Oxford University Press, Oxford (2012)
7. Kamshilin, A.A., Miridonov, S., Teplov, V., Saarenheimo, R., Nippolainen, E.: Photoplethysmographic imaging of high spatial resolution. Biomed. Opt. Express **2**(4), 996–1006 (2011)
8. McLeod, A.J., Baxter, J.S., Ameri, G., Ganapathy, S., Peters, T.M., Chen, E.C.: Detection and visualization of dural pulsation for spine needle interventions. Int. J. Comput. Assist. Radiol. Surg. **10**(6), 947–958 (2015)
9. McLeod, A.J., Baxter, J.S., de Ribaupierre, S., Peters, T.M.: Motion magnification for endoscopic surgery. In: SPIE Medical Imaging, vol. 9036, pp. 90360C–1–8 (2014)
10. Poh, M.Z., McDuff, D.J., Picard, R.W.: Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Trans. Biomed. Eng. **58**(1), 7–11 (2011)
11. Shao, D., Yang, Y., Liu, C., Tsow, F., Yu, H., Tao, N.: Noncontact monitoring breathing pattern, exhalation flow rate and pulse transit time. IEEE Trans. Biomed. Eng. **61**(11), 2760–2767 (2014)
12. Verkruysse, W., Svaasand, L.O., Nelson, J.S.: Remote plethysmographic imaging using ambient light. Opt. Express **16**(26), 21434–21445 (2008)