

Article

Bayesian-Inference Embedded Spline-Kerneled Chirplet Transform for Spectrum-Aware Motion Magnification

Enjian Cai ¹, Dongsheng Li ^{2,*}, Jianyuan Lin ² and Hongnan Li ³¹ Department of Civil Engineering, Tsinghua University, Beijing 100084, China; cej21@mails.tsinghua.edu.cn² Department of Civil and Environmental Engineering, Guangdong Engineering Center for Structure Safety and Health Monitoring, Shantou University, Shantou 515063, China; 19jylin@stu.edu.cn³ State Key Laboratory of Coastal & Offshore Engineering, Dalian University of Technology, Dalian 116023, China; hnli@dlut.edu.cn

* Correspondence: lids@stu.edu.cn

Abstract: The ability to discern subtle image changes over time is useful in applications such as product quality control, civil engineering structure evaluation, medical video analysis, music entertainment, and so on. However, tiny yet useful variations are often combined with large motions, which severely distorts current video amplification methods bounded by external constraints. This paper presents a novel use of spectra to make motion magnification robust to large movements. By exploiting spectra, artificial limitations and the magnification of small motions are avoided at similar frequency levels while ignoring large ones at distinct spectral pixels. To achieve this, this paper constructs spline-kerneled chirplet transform (SCT) into an empirical Bayesian paradigm that applies to the entire time series, giving powerful spectral resolution and robust performance to noise in nonstationary nonlinear signal analysis. The important advance reported is Bayesian-rule embedded SCT (BE-SCT); two numerical experiments show its superiority over current approaches. For applying to spectrum-aware motion magnification, an elaborate analytical framework is established that captures global motion, and use of the proposed BE-SCT for dynamic filtering enables a frequency-based motion isolation. Our approach is demonstrated on real-world and synthetic videos. This approach shows superior qualitative and quantitative results with less visual artifacts and more local details over the state-of-the-art methods.

Keywords: motion magnification; time-spectrum analysis; statistical inference; computer vision



Citation: Cai, E.; Li, D.; Lin, J.; Li, H. Bayesian-Inference Embedded Spline-Kerneled Chirplet Transform for Spectrum-Aware Motion Magnification. *Sensors* **2022**, *22*, 2794. <https://doi.org/10.3390/s22072794>

Academic Editor: Paolo Gastaldo

Received: 13 January 2022

Accepted: 18 March 2022

Published: 6 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Video motion magnification techniques have opened up a wealth of important applications. Examples include detecting a heartbeat vibration from tiny head motions [1] or blood flow [2], magnifying muscle tremors [3,4] to give an accurate clinical judgement, reconstructing speech information from small visual variations [5], evaluating material properties by the way it moves [6], estimating damage information of a building structure by measuring small vibrations in video [7–9], and lip reading [10]. However, some essential properties of dynamic objects become evident only when they move. For example, the muscles of an athlete when doing sports, the mechanical properties of a drone in flight, or the tremors of a Parkinson patient during walking. To open up new applications, this paper proposes to amplify variations at dynamic spectrum ranges in the entire time series, which makes motion amplification robust to occlusions and large motions.

Current video magnification techniques are classified into two categories: Lagrangian and Eulerian perspective. For Lagrangian approaches [11], tiny visual variations can be magnified by explicitly estimating feature points with optical flow, but corruptions affect amplification quality easily, since the local motion is represented by a single feature pixel point. The Eulerian techniques, on the other hand, do not estimate motions explicitly. Instead, they decompose video frames into representations to manipulate the local motions,

thus handling noise at deeper levels. In [2], an input video frame sequence is first decomposed into a multiscale stack (Laplacian or Gaussian pyramids); then, subtle changes are temporally filtered to find the variations to be amplified. When scaled and added back to the input images, a magnified output is rendered. With the complex-steerable pyramid [12–14], an input video frame sequence can be decomposed into a multi-orientation, multiscale stack [15]. Moreover, the use of phase-based motion processing has been considered not only in the context of motion magnification but also for many motion-related cases. In [5], phase decomposition is employed for extracting sound information from high-speed cameras, whereas in [16], the video phase information is applied to predict object material properties, and in [17], phase aids in estimating measurements of structural vibrations. In the context of motion amplification, the successful work in [15] extracts phase information through complex steerable filters and then magnifies it, the phase-based technique has better noise handling characteristics and supports larger amplification factors. In [18], a significant speedup without perceptual decrease in quality has been obtained; this work approximates the complex pyramid with the Riesz pyramid. While extremely successful for clean video signals, all these approaches assume that the objects of interest have very small motion—no large object motion or camera motion existing. Such large motions when processed result in large artifacts such as haloes or ripples, and our technique is specially designed to deal with these cases.

To deal with large motions, a layer-based video magnification approach was proposed in [19], with some help of a manually drawn mask by the user, outlining regions whose pixels are specified to be tracked and then magnified and yielding good magnification results. Whereas a mask indicates which pixels should be used, motion filter effects on the border of the mask cannot be ignored, leading to a certain spatial extent and eventually leaking across the mask edge. On the other hand, manual selection is time consuming and error prone; the selected region tracking is sensitive to occlusions and 3D object rotations. Furthermore, the alignment is based on a homography, which may generate wrong information for non-planar objects and a non-static camera. By using depth cameras and bilateral filters, the recent work in [20] proposed an alternative approach, making it possible for the amplification processing to be applied on pixels located at the same depth. In a sense, depth-aware motion processing extends the layer-based approach in replacing the manual selection mask by a weighted mask obtained from depth ranges, which avoids manual annotation to some extent. In addition, it also prevents the leaking problem in [19] by ignoring motion effects from different depth layers. However, this technique cannot cope with any moving objects; more importantly, the lack of depth knowledge will introduce inaccurate manual operations in processing. Based on the assumption that the large motion is typically linear at the scale of the small variations, the innovative processing framework [21] was proposed to magnify small deviations of linear motion by linking the response of a second-order Gaussian derivative to spatial acceleration. This work achieves impressive results for motion magnification; however, the downside is the inability to cope with nonlinear large motion. Inspired by the above approaches, this essay exploits time-frequency characteristics to automatically define the mask. In addition, based on the observation in [21,22], the significant differences are found in the frequency domain between these two kinds of variations, making our technique in principle suitable for large motion isolation.

Several techniques are available for time-frequency analysis, having played important roles in analyzing nonstationary signals. Among a number of analysis methods, short-time Fourier transform (STFT) [23–25], Wigner–Ville distribution (WVD) [26,27], wavelet transform (WT) [28,29], and Hilbert–Huang transform (HHT) [30,31] have been widely applied. For the STFT approach, since the STFT is based on traditional Fourier transform, the signal is assumed to be piecewise stationary at the scale of the window width, so showing weakness in accurate estimation for time-varying frequency. Whereas extremely successful for presenting excellent time-frequency representation for signals in terms of energy concentration, the WVD's bilinear structure creates the redundant cross terms that

cannot track the true time–frequency structure of the signal well, leading to the inaccurate estimation of instantaneous frequency. As another form of STFT method with an adjustable window size, the WT uses a large window for low-frequency components and a small window for high-frequency components, so it cannot achieve an accurate estimation for time-varying frequency as well [25]. Via applying the combination of empirical mode decomposition (EMD) and Hilbert spectral analysis, HHT offers a powerful way for non-stationary nonlinear signal analysis, which makes the instantaneous frequency meaningful. The need for spurious harmonics to represent nonlinear and nonstationary signals is also eliminated. However, the shortcomings of HHT include envelope fitting, mode mixing, end effects of EMD, and no uniform criterion for sifting stop, which may yield misleading results in nonstationary nonlinear signal analysis. By introducing an extra chirp kernel, which is characterized by the chirping rate parameter, the time–frequency atoms of the Chirplet transform (CT) can be sheared and shifted to match the signal in the time–frequency plane, thus showing superiority to WT and other time–frequency analysis approaches in analyzing nonstationary signals. Nonetheless, due to the inability of the chirp kernel to suit nonlinear-frequency-modulated (NLFM) signals, the CT fails the identification and extraction of the nonlinear frequency of the NLFM signal. By replacing the kernel of a frequency-shift operator and a frequency-rotate operator with spline kernel function [32], the spline-kernel chirplet transform (SCT) extends the capability of the CT and is able to produce a time–frequency representation with an excellent energy concentration for signals with nonlinearly time-varying instantaneous frequency, such that the instantaneous frequency of the NLFM signal can be accurately estimated. This paper begins with the SCT due to its superiority over other time–frequency analysis techniques. However, a critical shortcoming remains in current time–frequency methods, including SCT. These widely used techniques lack a statistical inference framework applicable to the entire time series [33–39]; their spectrum estimates on adjacent intervals cannot be formally related. Therefore, this paper adapts SCT by constructing a Bayesian statistical inference framework so it can be applied to wider practical projects.

In the following sections, the current parameterized time–frequency analysis techniques will be first discussed; then, we offer a statistical inference framework on how to model nonstationary time series on nonoverlapping intervals. The improved technique is experimentally evaluated by comparing against the state-of-the-art approaches in two numerical examples, and application of the proposed algorithm for video magnification in the presence of large motion is also shown to yield a superior performance over the existing amplification method.

2. Proposed Be-Sct

In this section, based on the traditional SCT (readers interested in this theory, please refer to [34,40–42]), the nonstationary time signal is modeled as a series of second-order stationary Gaussian processes defined on nonoverlapping spline function intervals. After SCT processing, a frequency domain random-walk model is utilized to relate the spectral representations of the Gaussian processes. The proposed algorithm efficiently calculates spectral updates by parallel complex Kalman filters; moreover, an expectation–maximization (EM) algorithm is utilized to estimate static and dynamic model parameters. The estimate is empirical Bayes, because it is computed conditional on the maximum likelihood parameter estimates.

SCT entails estimating the frequency content as a spline function of time for the non-stationary signal, and it is carried out by repeating spectrum estimation time intervals. Nevertheless, spectrum estimates on adjacent intervals are not regularly related. In contrast, current time–frequency methods are computationally intensive, achieve their high performance in signal-to-noise problems, and up to now have had limited application in practical time series analyses. Despite their usefulness for studying important problems, a critical shortcoming remains in current time–frequency methods including SCT: none of them offers an efficient statistical inference framework appropriate for the entire time series.

State-space modeling is an established, flexible inference framework for analyzing systems with properties that change over time [41,42]. In addition, this paradigm has been widely applied for the analysis of nonstationary time series with harmonic regression models [43], parametric time series models [44,45], and nonparametric time series models based on batch processing [46,47]. Therefore, on the basis of SCT, a plausible approach to analyze nonstationary and oscillatory time series can be proposed. By providing a flexible time-domain decomposition of the time series and a broadly applicable, empirical Bayes' framework for statistical inference, a comprehensive analysis framework for time-varying spectral analysis of nonstationary nonlinear time series can be achieved. The crucial advance reported is specially constructed Bayesian-rule embedded SCT (BE-SCT).

2.1. Theory

In the time-frequency model of BE-SCT, a nonstationary nonlinear time series observed can be defined as:

$$y_t = x_t + \varepsilon_t \quad (1)$$

where x_t is a second-order, zero mean, locally stationary Gaussian process, and ε_t is a zero mean, independent Gaussian noise with common variance σ_ε^2 . A common approach in the analysis of nonstationary time series is to assume a minimum interval length on which the data are stationary. The stationary intervals are indexed as: $i = 1, 2, \dots, I$, where I defines the number of distinct, nonoverlapping stationary intervals in x_t .

Based on the spectral representation theorem [48], the form of the observation model on stationary interval i is defined as:

$$\begin{aligned} Y_i &= X_i + \varepsilon_i \\ &= \Delta Z_i + \varepsilon_i \end{aligned} \quad (2)$$

where ε_i denotes an independent, zero mean Gaussian noise with common variance σ_ε^2 .

To relate the data on adjacent intervals, the Gaussian increment differences are assumed to be linked by the random walk model.

$$\Delta Z_i = \Delta Z_{i-1} + v_i \quad (3)$$

where v_i is assumed to be an independent, zero mean complex Gaussian process. In Equation (3), a stochastic continuity constraint is defined on the nonstationary time series in the frequency domain.

Followed by applying SCT, to represent the observation model Equation (2) in the frequency domain, the SCT operator is introduced, thus yielding the equation:

$$\begin{aligned} Y_{i,s}^{(SCT)} &= \Delta Z_{i,s}^{(SCT)} + \varepsilon_{i,s}^{(SCT)} \\ \Delta Z_{i,s}^{(SCT)} &= \Delta Z_{i-1,s}^{(SCT)} + v_{i,s}^{(SCT)} \end{aligned} \quad (4)$$

where s denotes the number of observations per stationary interval in the time-frequency plane, $Y_{i,s}^{(SCT)} = Y_i * SCT_s$, $\varepsilon_{i,s}^{(SCT)} = \varepsilon_i * SCT_s$ is a zero mean, complex Gaussian vector, and $v_{i,s}^{(SCT)} = v_i * SCT_s$ is also a zero mean, independent complex Gaussian vector. For ease of reading, the superscript " (SCT) " is omitted in the following derivations.

2.2. Algorithm

According to the linear complex Gaussian form of Equation (4), the sequence of increment differences [44] can be computed by a Kalman filter algorithm. The Gaussian increment difference estimates are assumed to have been computed on interval $(i-1)$, then for line s , a 1D complex Kalman filter algorithm for estimating $\Delta Z_{i,s}(\omega_s)$ on interval i can be obtained:

$$\begin{aligned}
\Delta Z_{i|i-1,s}(\omega_s) &= \Delta Z_{i-1|i-1,s}(\omega_s) \\
\sigma_{i|i-1,s}^2 &= \sigma_{i-1|i-1,s}^2 + \sigma_{v,s}^2 \\
\Delta Z_{i|i,s}(\omega_s) &= \Delta Z_{i|i-1,s}(\omega_s) + C_{i,s} \left(Y_{i,s} - \Delta Z_{i|i-1,s}(\omega_s) \right) \\
\sigma_{i|i,s}^2 &= (1 - C_{i,s}) \sigma_{i|i-1,s}^2
\end{aligned} \tag{5}$$

The Kalman gain for $i = 1, \dots, I$ and $s = 1, \dots, S$ can be computed as:

$$C_{i,s} = \left(\sigma_{\varepsilon}^2 + \sigma_{i|i-1,s}^2 \right)^{-1} \sigma_{i|i-1,s}^2 \tag{6}$$

The definition $i|u$ is the estimation on the stationary interval i based on all of the signal data observed through stationary interval u .

To efficiently analyze the functions of the increment differences at any time, the joint distribution of the increment differences in the time series can be computed using the fixed interval smoothing algorithm, which is defined as:

$$\begin{aligned}
\Delta Z_{i|I,s}(\omega_s) &= \Delta Z_{i|i,s}(\omega_s) + A_{i,s} \left(\Delta Z_{i+1|I,s}(\omega_s) - \Delta Z_{i+1|i,s}(\omega_s) \right) \\
\sigma_{i|I,s}^2 &= \sigma_{i|i,s}^2 + A_{i,s}^2 \left(\sigma_{i+1|I,s}^2 - \sigma_{i+1|i,s}^2 \right) \\
A_{i,s} &= \sigma_{i|i,s}^2 \left(\sigma_{i+1|i,s}^2 \right)^{-1}
\end{aligned} \tag{7}$$

In the smoothing algorithm, the initial conditions are $\Delta Z_{i,s}(\omega_s)$ and $\sigma_{i|I,s}^2$ for $i = I - 1, I - 2, \dots, 1$ and $s = 1, 2, \dots, S$. The covariance smoothing algorithm is used to obtain the covariances between any two states:

$$\sigma_{i,\mu|I,S} = A_{i,s} \sigma_{i+1,\mu|I,S} \tag{8}$$

where $1 \leq i \leq u \leq I$, Equations (7) and (8) are utilized to compute the joint distribution of the increment differences on all of the data. The distribution of any function of the increment differences can be computed by Monte Carlo methods [49,50], and a Monte Carlo estimate of its posterior probability density can be provided by the histogram of the function. The estimate process is empirical Bayes, since it is computed on the basis of maximum likelihood parameter estimates.

2.3. Model Parameters and Initial Condition Estimation

In the processing of Kalman filter (5) and (6), Kalman smoother (7), and covariance smoothing (8) algorithms, the initial state variances $\sigma_{0,s}^2$, the initial states $\Delta Z_{0,s}(\omega_s)$, and the model parameters $\sigma_{v,s}^2$ and σ_{ε}^2 are assumed to be known; then, an EM algorithm is used to obtain maximum likelihood estimates [46] of the parameters. The details are concluded as follows.

Firstly, the joint probability distribution of $\Delta Z_{1,s}(\omega_s)$ and $\Delta Y_{0,s}(\omega_s)$ at frequency s is expressed as:

$$\begin{aligned}
L_s &= p \left(\Delta Z_{0,s}(\omega_s) \mid \sigma_{v,s}^2 \right) \times \prod_{i=1}^I p \left(\Delta Z_{i,s}(\omega_s) \mid \Delta Z_{i-1,s}(\omega_s), \sigma_{v,s}^2 \right) \\
&\times \prod_{i=1}^I p \left(Y_{i,s}(\omega_s) \mid \Delta Z_{i,s}(\omega_s), \sigma_{\varepsilon}^2 \right)
\end{aligned} \tag{9}$$

Assume that the probability density of the initial state is obtained by:

$$p(\Delta Z_{0,s}(\omega_s)) = \left(\pi \sigma_{v,s}^2 \right)^{-1} \exp \left(- \frac{\| \Delta Z_{0,s}(\omega_s) \|^2}{\sigma_{v,s}^2} \right) \tag{10}$$

E-step: In iteration it of the E-step, on the basis of the previous estimates of the parameters and observed data from iteration $(it - 1)$, the expectation of the complete data log-likelihood can be calculated. For easy readability, Θ is introduced to represent the parameters $\{\sigma_\epsilon^2, \sigma_{v,s}^2, \Delta Z_{0,s}, \sigma_{0,s}^2\}$. Analyzing log and expectation to the likelihood yields:

$$E[\log L_s^{it} | \cdot Y_{1:L,S}, \Theta^{(it-1)}] = E \left[-\frac{\|\Delta Z_{0,S}^{it}(\omega_s)\|^2}{\sigma_{v,S}^{2,(it-1)}} - \frac{1}{\sigma_\epsilon^{2,(it-1)}} \sum_{i=1}^I \|Y_{i,S}^{it} - \Delta Z_{i,S}^{it}(\omega_s)\|^2 - (p-1) \log(\pi \sigma_\epsilon^{2,(it-1)}) - (p-1) \log(\pi \sigma_{v,S}^{2,(it-1)}) - \frac{1}{\sigma_{v,S}^{2,(it-1)}} \sum_{i=1}^I \|\Delta Z_{i,S}^{it}(\omega_s) - \Delta Z_{i-1,S}^{it}(\omega_s)\|^2 | Y_{1:L,S} \right] \tag{11}$$

Three quantities are required to be calculated to evaluate Equation (11).

$$\begin{aligned} \Delta Z_{i|L,S}^{it}(\omega_s) &= E[\Delta Z_{i,S}(\omega_s) | Y_{1:L,S}, \Theta^{(it-1)}] \\ W_{i|L,S}^{it} &= E[\|\Delta Z_{i,S}(\omega_s)\|^2 | Y_{1:L,S}, \Theta^{(it-1)}] \\ W_{i,i-1|L,S}^{it} &= E[\Delta Z_{i,S}(\omega_s) \Delta Z_{i-1,S}^* | Y_{1:L,S}, \Theta^{(it-1)}] \end{aligned} \tag{12}$$

By using the Kalman filter (5) and (6), Kalman smoothing (7), and covariance smoothing algorithms (8), these three quantities can be efficiently computed.

M-step: Let $\tau_{(v,s)}^{it} = 1 / \sigma_{v,s}^{2,it}$ and $\tau_\epsilon^{it} = 1 / \sigma_\epsilon^{2,it}$, then each gamma prior density is defined as:

$$e\rho(\tau | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} (\tau)^{\alpha-1} \exp(-\beta\tau) \tag{13}$$

For $\alpha > 1$ and $\beta > 0$, the expectation of log joint posterior density of each parameter is:

$$\begin{aligned} E[\log e\rho(\tau_\epsilon^{it}, \tau_{v,s}^{it}) Y_{1:p-1,s}, \Theta^{it-1}] &\propto \\ \log(e\rho(\tau_\epsilon^{it} | \alpha, \beta)) + \log(e\rho(\tau_{v,s}^{it} | \alpha, \beta)) &+ E[\log L_s^{it}] \end{aligned} \tag{14}$$

Equation (14) is required to be maximized with respect to $\tau_{v,s}^{it}$ and τ_ϵ^{it} , yielding the results:

$$\begin{aligned} \tau_{v,s}^{it} &= \frac{p-1+\alpha}{2 \sum_{i=1}^{it} (W_{i-1|L,S}^{it} - R\{W_{i,i-1|L,S}^{it}\}) + W_{1|L,S}^{it} + \beta} \\ \tau_\epsilon^{it} &= \frac{\alpha-1+S(p-1)}{\sum_{i,s} \left((Y_{i,S}^{it})^2 + W_{i|L,S}^{it} - 2R\{Y_{i,S}^{it} * \Delta Z_{i|L,S}^{it}\} \right) + \beta} \end{aligned} \tag{15}$$

In addition, each initial state $\Delta Z_{0,s}$ and initial variance $\sigma_{0,s}^2$ can be computed as:

$$\begin{aligned} \Delta Z_s(\omega_s) &= Y_1 * SCT_s \\ \sigma_s^2 &= \Delta Z_s \circ (\Delta Z_s)^* \end{aligned} \tag{16}$$

where \circ denotes the Hadamard product, the EM algorithm iterates between **E-steps** and **M-steps** until satisfying Equation (17) or $it = EM_{max}$, and EM_{max} is a predefined number of maximum iterations and $\epsilon \in (0, 0.001)$.

$$\frac{\|\Delta Z_{i|L,S}^{it} - \Delta Z_{i|L,S}^{it-1}\|^2}{\Delta Z_{i|L,S}^{it-1}} < \epsilon \tag{17}$$

The details are described in Algorithm 1.

Algorithm 1 BE-SCT

Equations

Input: matrix $Y_{i,s}$, maximum iteration of EM steps EM_{max} , minor integer ϵ .

while $it < EM_{max}$ and $\epsilon_{it} < \epsilon$ **do**

 Generating initial states $\Delta Z_{0,s}$ and initial variances $\sigma_{0,s}^2$ by Equation (16);

for $i = 2, \dots, I$ **do**

 Obtaining quantities in Equation (12) by recursively solving Equations (5) and (6) to evaluate Equation (11).

end for

for $i = I - 1, \dots, 1$ **do**

 Recursively computing Equations (7) and (8), to get quantities in Equations (11) and (12).

end for

 Obtaining final $\tau_{v,s}^{it}$ and τ_{ϵ}^{it} by Equation (15);

$it = it + 1$.

for $i = 2, \dots, I$ **do**

 On the basis of estimated parameters, recursively solving Equations (5) and (6) to perform the Kalman filter in matrix $Y_{i,s}$.

end for

end while

Output: Final matrix $Y_{i,s}^{(BE-SCT)}$ with s lines and i columns.

2.4. Numerical Experiments

In this section, two numerical simulations are used to demonstrate the effectiveness of the proposed BE-SCT. To add an extra degree of difficulty in analysis, additive Gaussian noise with a standard deviation of 0.1 and a mean of zero is artificially induced to the nonstationary nonlinear analytical signal. The sampling frequency is set to 100 Hz, and the time-frequency representation obtained by the proposed BE-SCT is compared with the continuous WT (CWT), HHT, and SCT.

The first example is given by:

$$f(t) = \begin{cases} \sin(2\pi(25t + 10 \sin t)) & (0 \leq t \leq 6 \text{ s}) \\ \sin(2\pi(34.2t)) & (6 \text{ s} \leq t \leq 10 \text{ s}) \end{cases} \quad (18)$$

The SNR of the first signal is 7.1203 dB, and the time-frequency representations generated by the CWT, HHT, SCT, and BE-SCT are shown in Figure 1. The wavelet is set to be 'cmor3-3', and the total scale is 256, the representation given by the CWT, as shown in Figure 1a. It can be seen that the CWT dissipates the energy around the instantaneous frequency at the high-frequency plane because of its coarse frequency resolution. Moreover, the representation is too blurry to reveal the time-frequency trajectory due to the high sensitivity to corruptions. In Figure 1b, HHT has higher anti-noise and robust performance compared with CWT; however, due to the misleading energy frequency distribution for corruptions and intrinsic modal information, the analytical consequence is too sparse to reveal the instantaneous frequency trajectory. As shown in Figure 1c, SCT provides excellent local estimates of signal features, but on account of its inability to offer a statistical inference framework appropriate for the entire time series, multiple instantaneous frequency trajectories are generated in the time-frequency plane. Figure 1d shows that the BE-SCT outperforms the CWT, the HHT, and the SCT as it clearly reveals the true time-frequency pattern of the analytical signal.

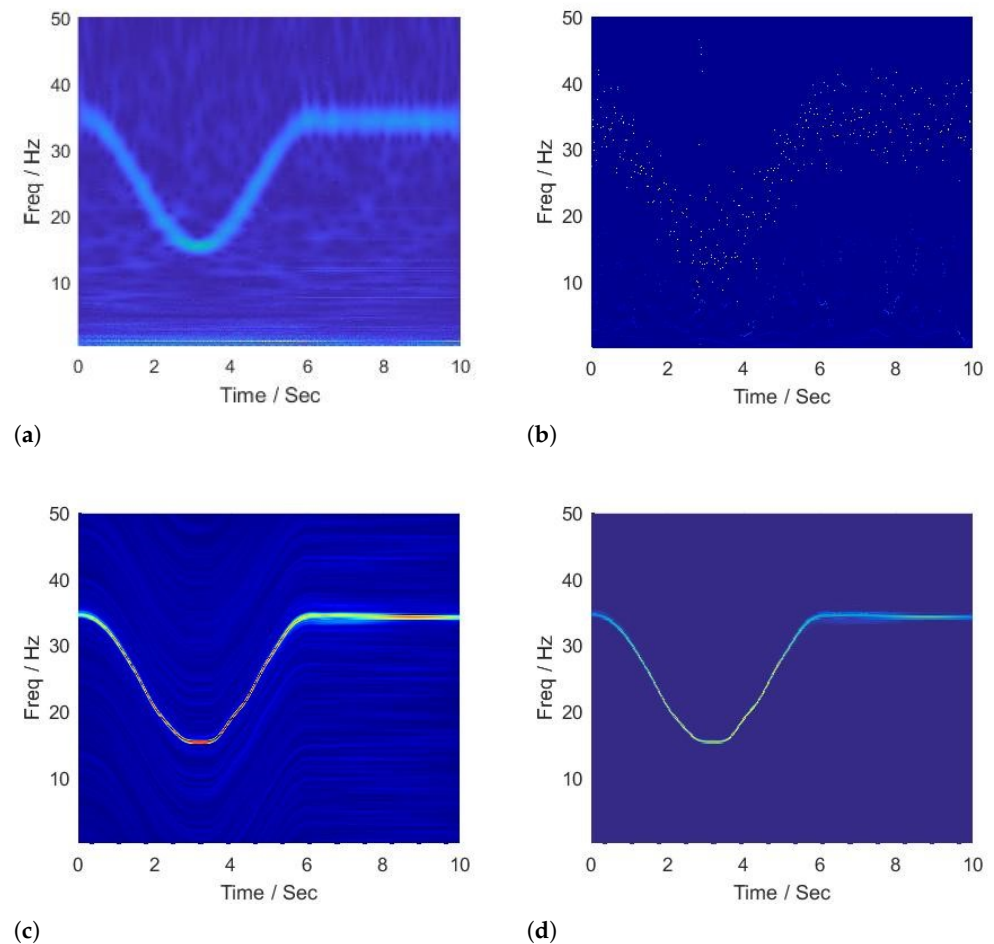


Figure 1. Time-frequency analysis of signal with instantaneous frequency in Equation (19) by (a) CWT, (b) HHT, (c) SCT, and (d) BE-SCT.

The second example is given by:

$$f(t) = \sin \left(30 + 50t + 60t^2 + 40 \sin t \right) \quad (0 \leq t \leq 10 \text{ s}) \quad (19)$$

The SNR of the second signal is 7.1177 dB, and the time-frequency representations generated by the CWT, HHT, SCT, and BE-SCT are shown in Figure 2. As shown in Figure 2a, CWT shows poor resolution in the time-frequency plane; besides, due to the reciprocal relationship between the center frequency of the wavelet function and window length, it is difficult for the representation given by the CWT to differentiate the true instantaneous frequency trajectory from the spurious frequency components introduced by the additive corruptions. The representation generated by the HHT is shown in Figure 2b, in which partial spurious frequency contents generated by nonstationary noise have been removed by EMD. However, some intrinsic modal information is mistaken for artifacts, so the true frequency trajectory is still hard to be identified by HHT. As shown in Figure 2c, the SCT offers the instantaneous frequency representation with an excellent concentration, and the most prominent trajectory can characterize the true time-varying frequency successfully. However, the spectrum estimates on adjacent spline functions are not formally related, resulting in the parallel spurious trajectories. On the other hand, as shown in Figure 2d, it is evident that the BE-SCT outperforms the CWT, the HHT, and the SCT; based on the precise parameters estimation, the adjacent spline functions are recursively linked, therefore giving its best performance in this high signal-to-noise spectrogram estimation problem.

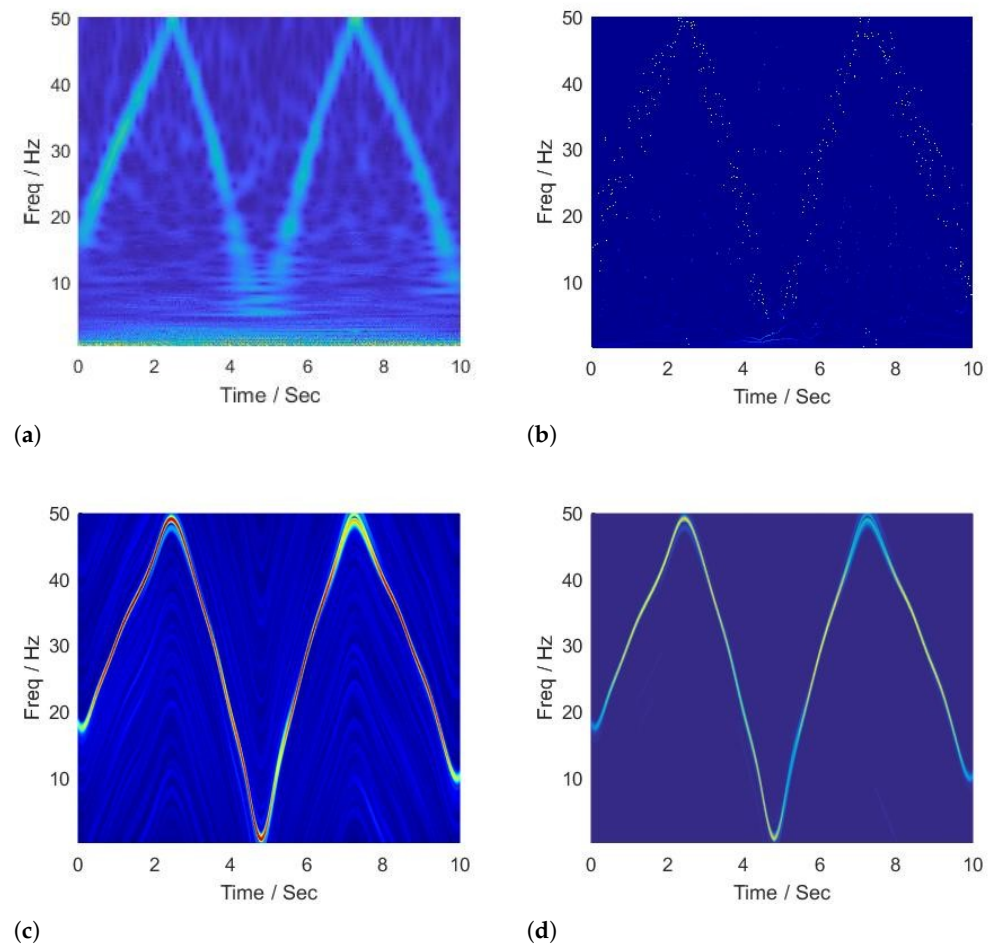


Figure 2. Time-frequency analysis of signal with instantaneous frequency in Equation (19) by (a) CWT, (b) HHT, (c) SCT, and (d) BE-SCT.

3. Spectrum-Aware Video Magnification

On the basis of BE-SCT, in this section, a spectrum-aware video magnification technique is presented to amplify small motions within large ones. Our technique has three main components:

1. On the basis of the earth mover's distance (EMOD) algorithm (readers interested in this theory, please refer to [22]), which avoids quantization and other binning problems, the moment function of original video motion information is temporally extracted;
2. By applying BE-SCT, the estimation stage seeks to understand the time-frequency characteristic of global nonstationary motions in analytical video;
3. With the appropriate prior knowledge, the dynamic ideal band-pass filter is used to remove large motions while preserving subtle ones.

Our proposed magnification pipeline is depicted in Figure 3.

3.1. Motion Metric Extraction

Carrying out the analysis of global motion information in video remains a challenging task due to the millions of pixels' respective temporal vibration signal existing in video. In engineering application areas and stereo-vision systems [51–57], the three-dimensional digital image correlation (3D DIC) and three-dimensional point tracking (3DPT) techniques are used to extract the full-field dynamic displacements of the analytical structure, and temporal vibration signals are further analyzed to obtain the material properties [58,59]. These methods are appropriate for structural engineering due to the law of mechanics;

however, limitations arise when analyzing an irregular video. The selection of degrees of freedom is time consuming and error prone. Inspired by recent research [60], which extracts the periodic pulsation of flame from the temporal image sequence based on the Euclidean distance and cross-correlation coefficient, EMOD is a method measuring a distance between two distributions. In this paper, every video frame is considered as a distribution, and EMOD is used to calculate the distance between each frame and the first frame. Therefore, temporal EMOD metrics can be conducted; then, BE-SCT is applied to these temporal metrics to generate the time-frequency estimation in the following parts. Detailed information of the EMOD can be seen in the reference [61].

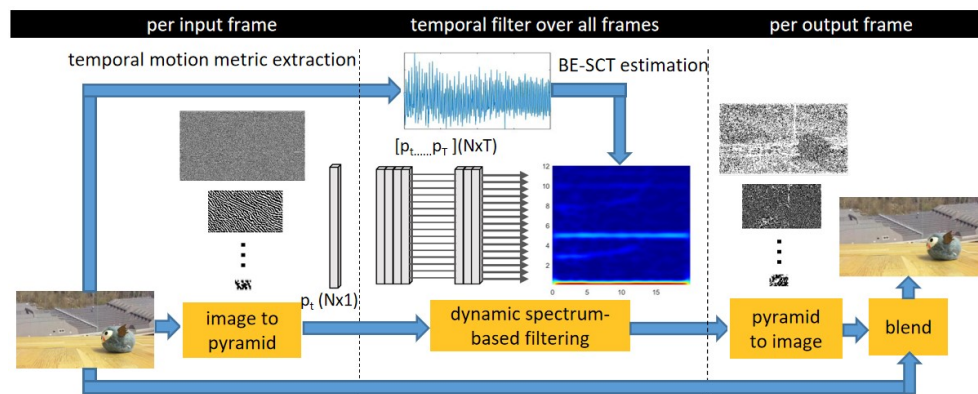


Figure 3. Video spectrum-aware magnification pipeline. Our approach does not require manual region annotation nor additional depth information as done in conventional techniques; instead, by employing the proposed BE-SCT, the intrinsic frequency characteristics can be understood to achieve the goal of adaptive large motions isolation, meanwhile avoiding the nonlinear limitation in the Eulerian acceleration approach.

3.2. Dynamic Spectrum-Aware Filtering

Applying the specially constructed BE-SCT, which has been verified for the superiority over other time-frequency estimation algorithms, the true time-varying frequency pattern of global motions in analytical video can be precisely obtained. Hereby, it can be observed that at the scale of subtle visual changes, the frequency of large motion is relatively small. By only magnifying small deviations of dynamically selected frequency ranges, this method arrives at temporal spectrum-aware filtering magnification.

For easy readability of the temporal spectrum-aware filtering, a time-domain mathematics model is established for illustration. Consider the time domain of intensity changes denoted by $I(x, y, t)$ at position (x, y) and time t based on the significant anti-noise performance of complex steerable pyramid, small temporal variations in the spatial offset of edges can be converted to subtle temporal changes in polar coordinates of the complex filter responses in the pyramid. Therefore, in the temporal mathematics model, the temporal variations are first turned into the frequency domain by Fourier transform, which is $S(\omega, \rho, t)$. Then, based on the observation that the large motions differ evidently from small ones in frequency property, the sophisticated time domain of intensity variations is reconstructed in the frequency domain as a combination of two components:

$$S(\omega, \rho, t) = S_{\rho a}(\omega, \rho, t) + S_{\rho b}(\omega, \rho, t) \quad (20)$$

where $S_{\rho a}(\omega, \rho, t)$ denotes the variations component with spectral amplitude above the threshold ρ , and $S_{\rho b}(\omega, \rho, t)$ stands for the opposite.

Ultimately, it is crucial to develop a self-adapting estimation for the spectrum-aware filtering to separate these two components. Returning to the proposed BE-SCT, it can act not only as the “detector” of the property of global video motions but also adaptively isolate the large motions from the small ones as deeper utilized. On the basis of the time-varying spectrogram by BE-SCT, this method further constructs a dynamic frequency-based

filtering algorithm to handle the challenging isolation task. In the actual operation, similar to the common ideal band-pass filter, the time-domain dynamic weighting function of the spectrum-aware filtering is defined as:

$$W_{\omega,\rho,t} = \begin{cases} 1, & \rho \in [\rho_l, \rho_h] \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

where $[\rho_l, \rho_h]$ stands for the identified amplitude-frequency ranges in the time-domain spectrogram generated by BE-SCT. ρ_l is the minimum amplitude bound, which is not critical, since small noise can be negligible after arbitrary simple built-in spatial filtering algorithms. ρ_h is the maximum amplitude threshold used for eliminating the large motions, which can be experience-modifiable.

4. Experimental Results

To verify the effectiveness of the proposed spectrum-aware approach, experiments on real sequences as well as on a synthetically generated one with ground truth are performed. This paper only assesses the real videos' performance qualitatively, whereas for the synthetic sequence, the quantitative evaluation is taken against ground truth. For all videos, the video frames are processed in a YIQ color space. In the contrast tests, a complex steerable pyramid with octave bandwidth filters and four orientations is used to decompose each frame into phase and magnitude. The results demonstrate that state-of-the-art techniques optimized for subtle variations generate blurs and artifacts when handling large motions. Our technique fully utilizes the powerful BE-SCT, significantly reduces haloes or corruptions, and increases the scope of its applicability.

4.1. Real-Life Sequences

Figure 4 shows a cat toy moving on the table, which coincided with the high-frequency vibration perpendicular to the circle trajectory. The goal of this experiment is to magnify the vibration with amplification factor $\alpha = 8$; the motion above the black arrow is recorded in the spatio-temporal slice indicated with the green line over the raw video. The phase-based motion magnification generates substantial artifacts due to the large movement on the table. The Eulerian-acceleration approach relies on the second-order filter; therefore, the nonlinear motions in the background are magnified while inducing tiny blurring effects, as seen in the figure. Our proposed technique manages to achieve this by amplifying the variations at the pixels that lie in the time-varying frequency property estimated by BE-SCT, thus magnifying the vibration of the toy and separating the motion along the trajectory on the table.

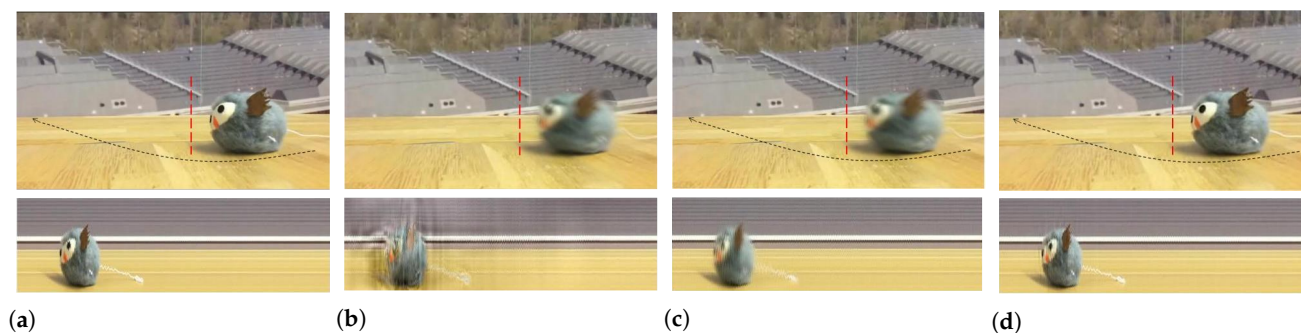


Figure 4. A cat toy vibrating at a high frequency, along with the large amplitude movement of a circle trajectory depicted by the black arrow. Four frames indicating the toy's trajectory are shown in each top row, the bottom rows show the spatio-temporal line corresponding to the green line in the relevant video frames. (a) Original video. (b) Phase-based video magnification. (c) Eulerian-acceleration magnification. (d) Our proposed spectrum-aware magnification. The proposed magnification approach can clearly reveal the vibration of the cat toy without inducing blurs and artifacts [21].

Figure 5 demonstrates various motion amplification results for a gun shooting video with magnification factor $\alpha = 8$. In this case, the recoil of the gun induces subtle movement in the arm muscles. To preform an in-depth and meticulous analysis, the movements of the bracelet, upper limb, and the forearm are recorded in the spatio-temporal slices indicated with three green lines over the original sequence. Due to the strong arm movement, the phase-based processing induces ripples and motion artifacts, which cover the subtle motion in the muscles. The Eulerian-acceleration method only magnifies the nonlinear motion, leading to the loss of linear subtle movement. Our proposed technique not only magnifies the intensity changes of the arm muscles but also magnifies clearly the intensity variations of the bracelet, which is caused by the reflection of the muscles, as shown in the plot on the bottom-right of Figure 5.

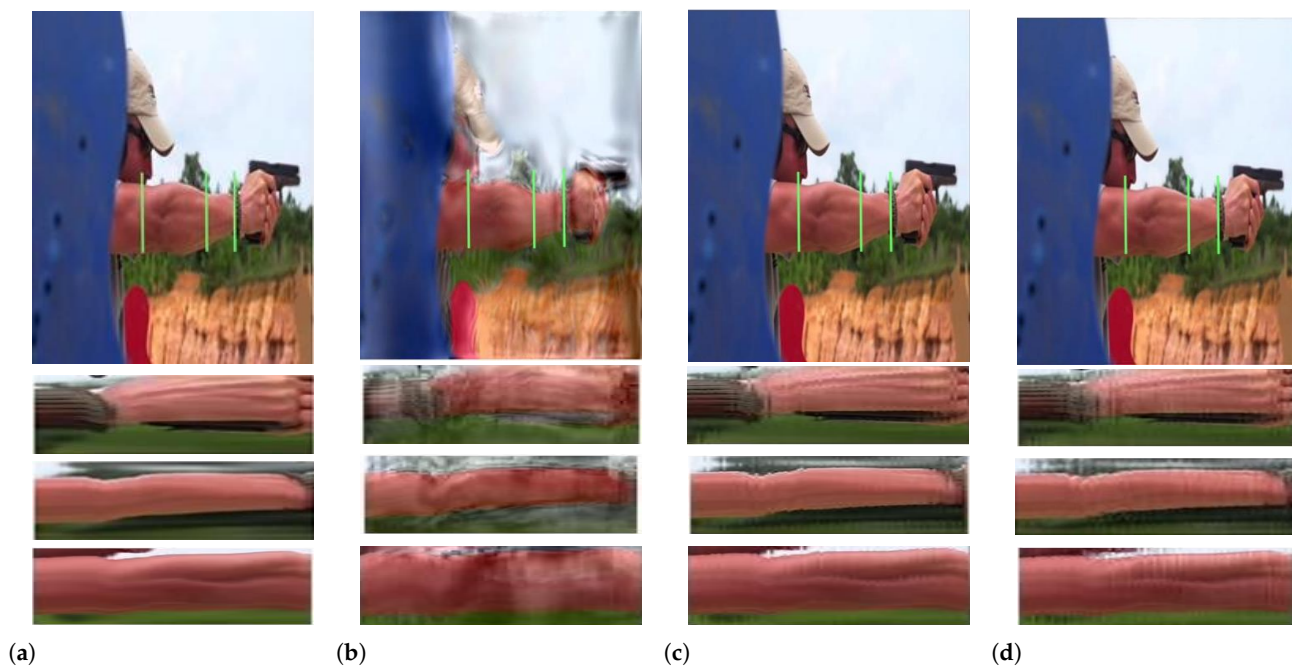


Figure 5. In the gun shooting sequence, the strong recoil causes small vibrations of the arm. The spatio-temporal slice is shown at different positions with three green lines over the sequence for each processing. (a) Original video frame. (b) Phase-based video magnification. (c) Eulerian-acceleration magnification. (d) Our proposed spectrum-aware magnification. The Eulerian-acceleration approach only magnifies the nonlinear motion by linking the response of a second-order Gaussian derivative, whereas the phase-based method results in large blurs and artifacts. Our proposed method magnifies the arm movements correctly without being affected by the background clutter [21].

In Figure 6, the picture shows a transparent bottle with water being pulled sideways on the smooth surface, whereas the level of water in the bottle fluctuates sharply, as shown in the original sequence. The magnification factor $\alpha = 8$ is chosen for each video processing. According to the contrast experiments, the phase-based approach generates significant blurring artifacts caused by the bottle moving. On the other hand, similar but more precise than Eulerian-acceleration processing, our approach is able to correctly amplify the desired motion—oscillation of the water level, while not inducing substantial blurring artifacts.

Figure 7 shows magnification results for iris wobbling, combined with large-scale eye horizontal movements, and sets the magnification factor $\alpha = 15$. As demonstrated in the figure (top-right), when applied to the video with the phase-based technique, the small motion remains hard to be seen because it is overshadowed by the then-magnified large motions and its blurring artifacts. Our spectrum-aware magnification maintains the local motions of the iris wobbling. Eulerian acceleration does magnify segmental temporal variations; however, it kills more useful information than our approach.

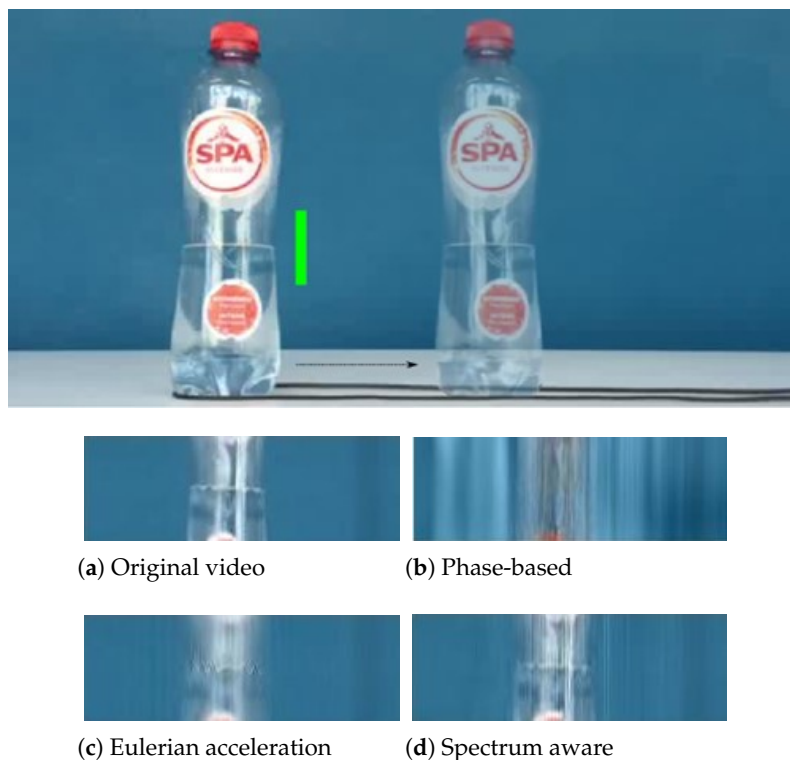


Figure 6. The water oscillating in a bottle while the bottle is being pulled sideways on a smooth surface. The green stripe indicates the locations at which the dynamic movements are temporally detected from the video. Compared to the state-of-the-art approaches, our proposed magnification method is able to amplify the oscillations in the water while not inducing substantial blurs [21].

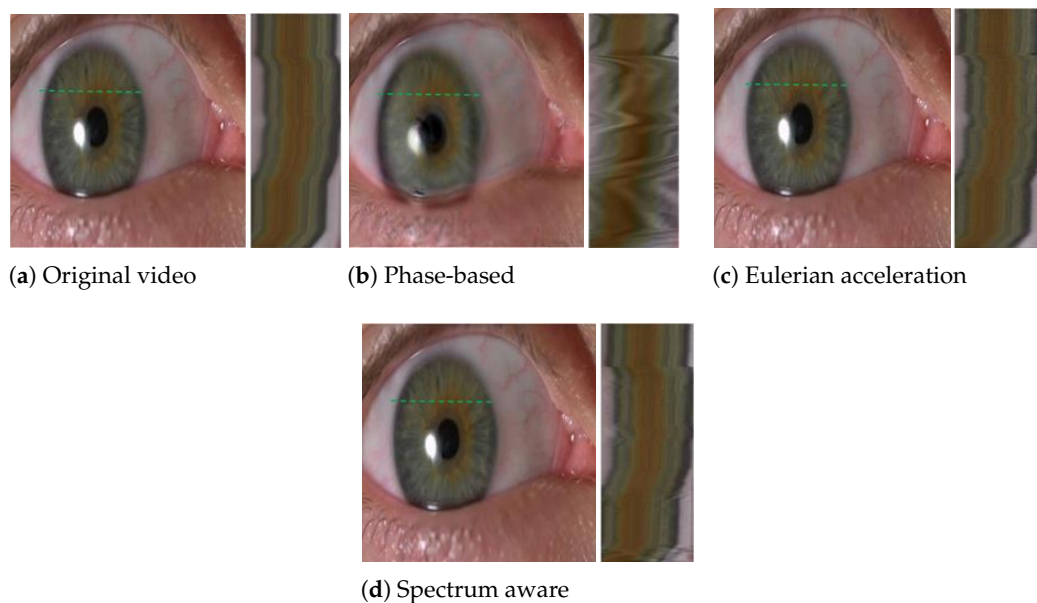


Figure 7. The eye video and its magnification with the phase-based approach, the Eulerian-acceleration approach, and our spectrum-aware processing. The spatio-temporal slice is shown in each approach for the green stripe (top-left). This video demonstrates an eye moving along the horizontal direction, as shown in the original sequence; such wobbling is too subtle to be observed (top-left). The global motion of the eye generates significant blurring artifacts when processed with the phase-based approach. However, processing the sequence with Eulerian acceleration and our approach show clearly that the iris wobbles as the eye moves; through the in-depth comparison, more local details can be preserved in our approach [19].

To quantitatively evaluate the performance of the proposed method compared with traditional methods, commonly used objective metrics including peak signal-to-noise ratio (PSNR) and mean absolute error (MAE) are further introduced, which are measured over the whole image in all frames. The range of MAE at the numerical span is $[0, 1]$; the closer to 0, the more similarities there are in the two images. Meanwhile, PSNR follows the opposite monotonicity rule. Results for all the real-life videos are given in Figures 8–11, respectively. It can be seen that the proposed method achieves higher values of PSNR and lower values of MAE than the traditional methods, quantitatively validating its superiority in terms of magnifying subtle changes and achieving the best anti-noise results.

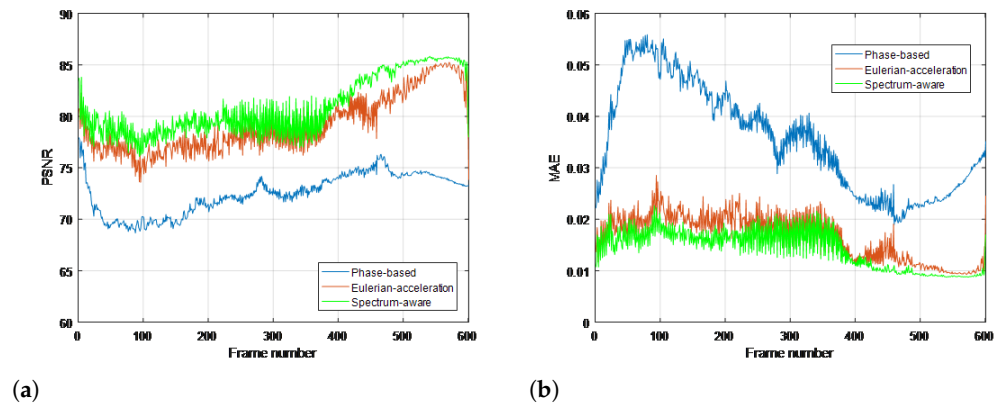


Figure 8. Objective metrics with ground truth for each magnification processing using (a) PSNR and (b) MAE in cat toy video

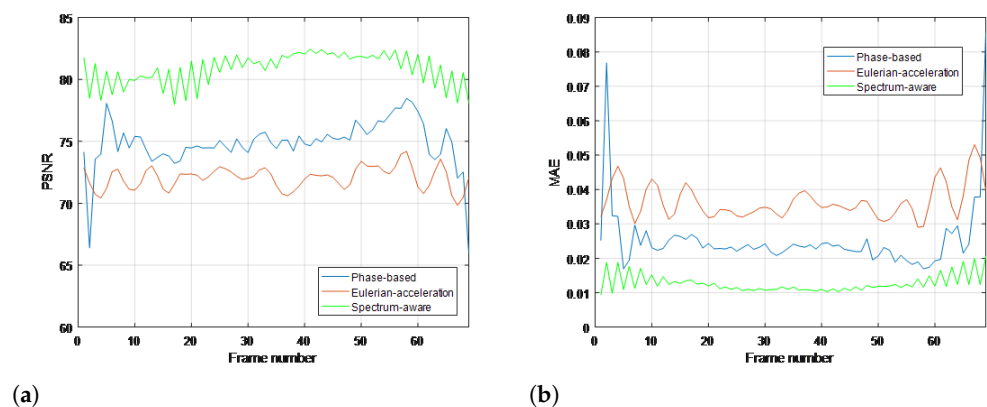


Figure 9. Objective metrics with ground truth for each magnification processing using (a) PSNR and (b) MAE in gun shooting video.

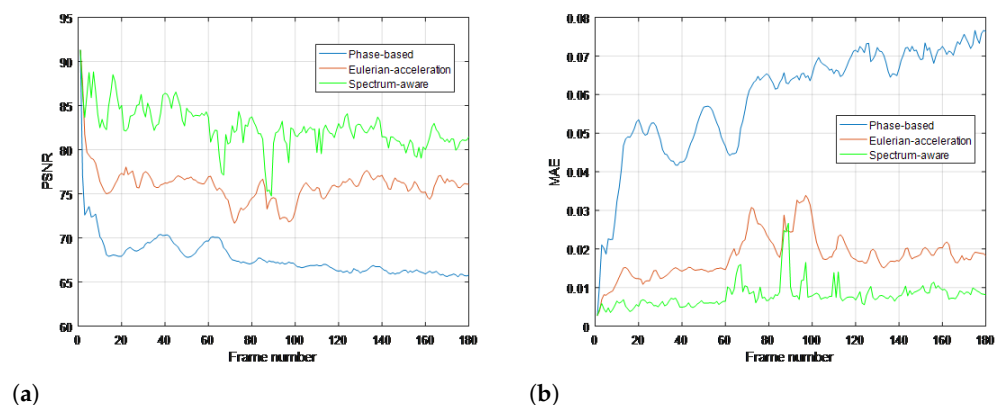


Figure 10. Objective metrics with ground truth for each magnification processing using (a) PSNR and (b) MAE in bottle video.

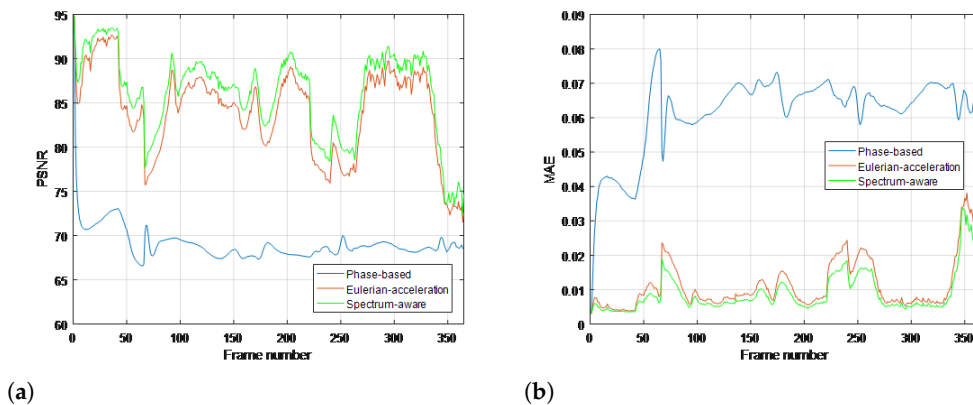


Figure 11. Objective metrics with ground truth for each magnification processing using (a) PSNR and (b) MAE in eye video.

4.2. Synthetic Sequence

In Figure 12, the picture demonstrates a synthetic ball that moves horizontally on the screen from the left to right corner; the radius of the ball is set to 10 pixels, and the velocity of movement is 1 pixel/frame. The ball vibrations are modeled in the vertical direction as a sine wave, with a maximum value of 1 pixel. The vibration frequency is 3 cycle/sec, and the frame rate is set to 30 frame/sec. For ground truth amplification, the temporal changes are magnified by two times without changing any other parameters. A complex steerable pyramid is applied for all contrast processes, with octave bandwidth filters and four orientations, which only amplify the pyramid level with a magnification factor of 5.



Figure 12. Synthetic video. A ball moves from left to right along with a tiny vibration in the vertical direction [21].

Objective results are given in Figure 13. Statistically speaking, our proposed algorithm yields the best performance for deriving the most significant fidelity. However, on the other hand, due to the innovative construction of second-order Gaussian derivative, the magnified sequence processed by the Eulerian-acceleration approach shows certain regularity in the time domain, which provides us some beneficial enlightenment of fundamental improvements for the future work.

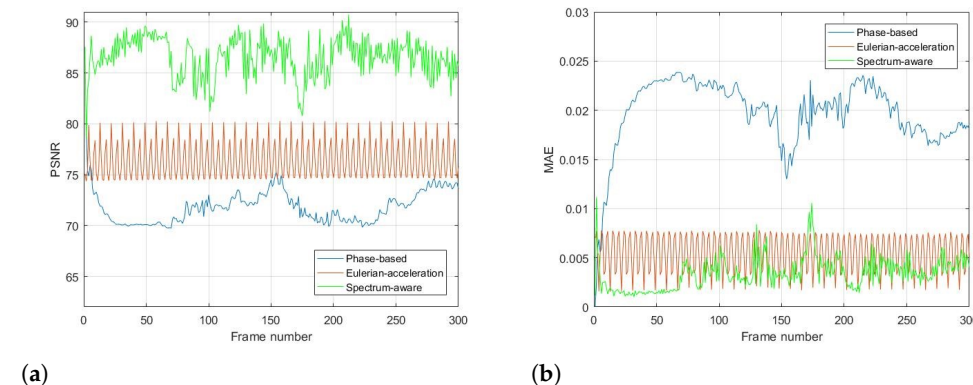


Figure 13. Objective metrics with ground truth for each magnification processing using (a) PSNR and (b) MAE in synthetic video.

5. Discussion

Limitation of our approach. As demonstrated in the controlled experiments for synthetic sequence magnification, compared with the Eulerian-acceleration approach, our magnified results show relatively irregular temporal variations. However, our algorithm achieves the best anti-noise performance and retains the most details over other methods, as verified in each controlled experiment. If required to recover regular small motions from large ones as possible in some special project applications, the Eulerian-acceleration approach may be the best choice in spite of its other flaws. Thus, to establish a more comprehensive analytical framework, some powerful shift rules will be constructed in the pyramid level in future research.

Performance superiority over state of the art. The superiority of our work can not only be summarized in the context of time-frequency analysis but also in video motion further revelation. For the time-varying spectral analysis of nonstationary nonlinear signals, this paper introduces the SCT, which has been validated for its excellent local estimates of data features. A statistical inference framework is established to efficiently relate its spectral estimates across local intervals; the important advance that this paper presents is BE-SCT. In Figures 1 and 2, BE-SCT depicts clearly the time-varying spectrum trajectory of two nonstationary nonlinear signals, whereas the state-of-the-art approaches do not. For large movements isolation, this method sidesteps the problem of Eulerian-acceleration processing for being easily affected by the nonlinear background clutter. The results in Figure 4–13 for processing whether real-life or synthetic sequences show the robustness superiority of our spectrum-aware technique. Therefore, in the future work, further extensions will be presented of its anti-noise ability.

6. Conclusions

Standard video magnification techniques cannot reliably handle large motions, which are bounded by excessive user annotations, additional depth information, their inability to operate nonlinear background clutter, and so on. By exploiting the spectrum characteristic of global motions in analytical video, we are not restricted by such limitations and can magnify unconstrained videos.

To construct a powerful time-varying spectral analytical framework, the spectral representation theorem-based inference model is adapted to SCT. Then, with the assistance of EMOD extraction, background large movement values are ignored by filter responses at spectrum layers.

Spectrum-aware motion magnification is demonstrated on several real-world and synthetic sequences. We show that our approach performs well, has better anti-noise characteristics, and has less background edge artifact than the state of the art. Improving robustness in the pyramid level so that it works at higher magnification is an important direction for future work.

Author Contributions: Conceptualization and methodology: D.L.; methodology and validation: D.L. and E.C.; formal analysis: E.C. and J.L.; resources, data curation and writing—original draft preparation: E.C.; writing—review and editing: D.L., H.L. and J.L.; visualization and supervision: D.L.; project administration and funding acquisition: D.L. and H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research work is jointly supported by the National Natural Science Foundation of China (51778103 and 52078284), Guangdong Basic and Applied Basic Research Foundation (2021A1515011770), STU Scientific Research Foundation for Talents (NTF18012), and Open Projects Foundation (No. BHSKL20-10-KF) of State Key Laboratory for Health and Safety of Bridge Structures.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Balakrishnan, G.; Durand, F.; Guttag, J. Detecting Pulse from Head Motions in Video. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 3430–3437.
2. Wu, H.Y.; Rubinstein, M.; Shih, E.; Guttag, J.; Durand, F.; Freeman, W. Eulerian Video Magnification for Revealing Subtle Changes in the World. *ACM Trans. Graph. (TOG)* **2012**, *31*, 1–8. [[CrossRef](#)]
3. Aziz, N.A.; Tannemaat, M.R. A Microscope for Subtle Movements in Clinical Neurology. *Neurology* **2015**, *85*, 920–920. [[CrossRef](#)] [[PubMed](#)]
4. Wang, Y.; See, J.; Oh, Y.H.; Phan, R.C.W.; Rahulamathavan, Y.; Ling, H.C.; Tan, S.W.; Li, X. Effective recognition of facial micro-expressions with video motion magnification. *Multimed. Tools Appl.* **2017**, *76*, 21665–21690. [[CrossRef](#)]
5. Davis, A.; Rubinstein, M.; Wadhwa, N.; Mysore, G.J.; Durand, F.; Freeman, W.T. The Visual Microphone: Passive Recovery of Sound from Video; *Assoc. Comput. Mach. (ACM)* **2014**, *33*, 79. [[CrossRef](#)]
6. Dorn, C.J.; Mancini, T.D.; Talken, Z.R.; Yang, Y.; Kenyon, G.; Farrar, C.; Mascareñas, D. Automated Extraction of Mode Shapes Using Motion Magnified Video and Blind Source Separation. In *Topics in Modal Analysis & Testing*; Springer: Berlin/ Heidelberg, Germany, 2016; Volume 10, pp. 355–360.
7. Cha, Y.J.; Chen, J.G.; Büyükoztürk, O. Output-Only Computer Vision Based Damage Detection Using Phase-Based Optical Flow and Unscented Kalman Filters. *Eng. Struct.* **2017**, *132*, 300–313. [[CrossRef](#)]
8. Sarrafi, A.; Mao, Z.; Niezrecki, C.; Poozesh, P. Vibration-Based Damage Detection in Wind Turbine Blades using Phase-Based Motion Estimation and Motion Magnification. *J. Sound Vib.* **2018**, *421*, 300–318. [[CrossRef](#)]
9. Fazio, N.L.; Leo, M.; Perrotti, M.; Lollino, P. Analysis of the Displacement Field of Soft Rock Samples During UCS Tests by Means of a Computer Vision Technique. *Rock Mech. Rock Eng.* **2019**, *52*, 3609–3626. [[CrossRef](#)]
10. Zhang, X.; Sheng, C.; Liu, L. Lip Motion Magnification Network for Lip Reading. In Proceedings of the 2021 7th International Conference on Big Data and Information Analytics (BigDIA), Chongqing, China, 23–31 October 2021; pp. 274–279. [[CrossRef](#)]
11. Liu, C.; Torralba, A.; Freeman, W.T.; Durand, F.; Adelson, E.H. Motion Magnification. *ACM Trans. Graph. (TOG)* **2005**, *24*, 519–526. [[CrossRef](#)]
12. Freeman, W.T.; Adelson, E.H. The Design and Use of Steerable Filters. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 891–906. [[CrossRef](#)]
13. Portilla, J.; Simoncelli, E.P. A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients. *Int. J. Comput. Vis.* **2000**, *40*, 49–70. [[CrossRef](#)]
14. Simoncelli, E.P.; Freeman, W.T. The Steerable Pyramid: A Flexible Architecture for Multi-Scale Derivative Computation. In Proceedings of the International Conference on Image Processing, Washington, DC, USA, 23–26 October 1995; pp. 444–447.
15. Wadhwa, N.; Rubinstein, M.; Durand, F.; Freeman, W.T. Phase-Based Video Motion Processing. *ACM Trans. Graph. (TOG)* **2013**, *32*, 1–10. [[CrossRef](#)]
16. Davis, A.; Bouman, K.L.; Chen, J.G.; Rubinstein, M.; Durand, F.; Freeman, W.T. Visual Vibrometry: Estimating Material Properties from Small Motion in Video. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 5335–5343.
17. Chen, J.G.; Wadhwa, N.; Cha, Y.J.; Durand, F.; Freeman, W.T.; Buyukozturk, O. Modal Identification of Simple Structures with High-Speed Video Using Motion Magnification. *J. Sound Vib.* **2015**, *345*, 58–71. [[CrossRef](#)]
18. Wadhwa, N.; Rubinstein, M.; Durand, F.; Freeman, W.T. Riesz Pyramids for Fast Phase-Based Video Magnification. In Proceedings of the 2014 IEEE International Conference on Computational Photography (ICCP), Santa Clara, CA, USA, 2–4 May 2014; pp. 1–10.
19. Elgharib, M.; Hefeeda, M.; Durand, F.; Freeman, W.T. Video Magnification in Presence of Large Motions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 4119–4127.
20. Kooij, J.F.; van Gemert, J.C. Depth-Aware Motion Magnification. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 467–482.
21. Zhang, Y.; Pintea, S.L.; Van Gemert, J.C. Video Acceleration Magnification. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 529–537.
22. Wu, X.; Yang, X.; Jin, J.; Yang, Z. Amplitude-Based Filtering for Video Magnification in Presence of Large Motion. *Sensors* **2018**, *18*, 2312. [[CrossRef](#)]
23. Nawab, S.H. Short-Time Fourier Transform. In *Advanced Topics in Signal Processing*; Prentice-Hall, Inc.: Upper Saddle River, NY, USA, 1988.
24. Kwok, H.K.; Jones, D.L. Improved Instantaneous Frequency Estimation Using an Adaptive Short-Time Fourier Transform. *IEEE Trans. Signal Process.* **2000**, *48*, 2964–2972. [[CrossRef](#)]
25. Kaiser, G. *A Friendly Guide to Wavelets*; Birkhäuser: Boston, MA, USA, 2010.
26. Blodt, M.; Bonacci, D.; Regnier, J.; Chabert, M.; Faucher, J. On-Line Monitoring of Mechanical Faults in Variable-Speed Induction Motor Drives Using the Wigner Distribution. *IEEE Trans. Ind. Electron.* **2008**, *55*, 522–533. [[CrossRef](#)]
27. Rosero, J.A.; Romeral, L.; Ortega, J.A.; Rosero, E. Short-Circuit Detection by Means of Empirical Mode Decomposition and Wigner–Ville Distribution for PMSM Running Under Dynamic Condition. *IEEE Trans. Ind. Electron.* **2009**, *56*, 4534–4547. [[CrossRef](#)]

28. Coppola, L.; Liu, Q.; Buso, S.; Boroyevich, D.; Bell, A. Wavelet Transform as an Alternative to the Short-Time Fourier Transform for the Study of Conducted Noise in Power Electronics. *IEEE Trans. Ind. Electron.* **2008**, *55*, 880–887. [[CrossRef](#)]
29. Bouzida, A.; Touhami, O.; Ibtouen, R.; Belouchrani, A.; Fadel, M.; Rezzoug, A. Fault Diagnosis in Industrial Induction Machines through Discrete Wavelet Transform. *IEEE Trans. Ind. Electron.* **2010**, *58*, 4385–4395. [[CrossRef](#)]
30. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The Empirical Mode Decomposition and the Hilbert Spectrum for Non-linear and Non-Stationary Time Series Analysis. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
31. Huang, N.E.; Wu, Z. A Review on Hilbert-Huang Transform: Method and its Applications to Geophysical Studies. *Rev. Geophys.* **2008**, *46*. [[CrossRef](#)]
32. Yang, Y.; Peng, Z.; Meng, G.; Zhang, W. Spline-Kernelled Chirplet Transform for the Analysis of Signals with Time-Varying Frequency and its application. *IEEE Trans. Ind. Electron.* **2011**, *59*, 1612–1621. [[CrossRef](#)]
33. Chatterji, S.; Blackburn, L.; Martin, G.; Katsavounidis, E. Multi-resolution Techniques for the Detection of Gravitational-Wave Bursts. *Class. Quantum Gravity* **2004**, *21*, 1809–1818. [[CrossRef](#)]
34. Haigh, J.D.; Winning, A.R.; Toumi, R.; Harder, J.W. An Influence of Solar Spectral Variations on Radiative Forcing of Climate. *Nature* **2010**, *467*, 696–699. [[CrossRef](#)]
35. Zheng, X.; Chen, B.M. *Stock Market Modeling and Forecasting*; Springer: London, UK, 2013.
36. Truccolo, W.; Eden, U.T.; Fellows, M.R.; Donoghue, J.P.; Brown, E.N. A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. *J. Neurophysiol.* **2005**, *93*, 1074–1089. [[CrossRef](#)]
37. Mitra, P. *Observed Brain Dynamics*; Oxford University Press: Oxford, UK, 2007.
38. Yilmaz, T.; Foster, R.; Hao, Y. Detecting Vital Signs with Wearable Wireless Sensors. *Sensors* **2010**, *10*, 10837–10862. [[CrossRef](#)]
39. Quatieri, T.F. *Discrete-Time Speech Signal Processing: Principles and Practice*; Pearson Education India: Delhi, India, 2006.
40. Unser, M.; Aldroubi, A.; Eden, M. B-Spline Signal Processing. II. Efficiency Design and Applications. *IEEE Trans. Signal Process.* **1993**, *41*, 834–848. [[CrossRef](#)]
41. Fahrmeir, L.; Tutz, G.; Hennevogl, W.; Salem, E. *Multivariate Statistical Modelling Based on Generalized Linear Models*; Springer: New York, NY, USA, 1994.
42. Durbin, J.; Koopman, S.J. *Time Series Analysis by State Space Methods*; Oxford University Press: Oxford, UK, 2012.
43. Qi, Y.; Minka, T.P.; Picara, R.W. Bayesian Spectrum Estimation of Unevenly Sampled Nonstationary Data. In Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, USA, 13–17 May 2002.
44. Kitagawa, G.; Gersch, W. *Smoothness Priors Analysis of Time Series*; Springer Science & Business Media: Heidelberg, Germany, 1996.
45. Tarvainen, M.P.; Hiltunen, J.K.; Ranta-Aho, P.O.; Karjalainen, P.A. Estimation of Nonstationary EEG with Kalman Smoother Approach: An Application to Event-Related Synchronization (ERS). *IEEE Trans. Biomed. Eng.* **2004**, *51*, 516–524. [[CrossRef](#)]
46. Ba, D.; Babadi, B.; Purdon, P.L.; Brown, E.N. Robust Spectrotemporal Decomposition by Iteratively Reweighted Least Squares. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5336–5345. [[CrossRef](#)]
47. Kim, S.E.; Behr, M.K.; Ba, D.; Brown, E.N. State-Space Multitaper Time-Frequency Analysis. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 5–14. [[CrossRef](#)]
48. Shumway, R.H.; Stoffer, D.S.; Stoffer, D.S. *Time Series Analysis and Its Applications*; Springer: New York, NY, USA, 2000.
49. Doucet, A.; De Freitas, N.; Gordon, N. An Introduction to Sequential Monte Carlo Methods. In *Sequential Monte Carlo Methods in Practice*; Springer: New York, NY, USA, 2001; pp. 3–14.
50. Carlin, B.P.; Louis, T.A. *Bayesian Methods for Data Analysis*; CRC Press: Boca Raton, FL, USA, 2008.
51. Bebernis, T.J.; Ehrhardt, D.A. High-Speed 3D Digital Image Correlation Vibration Measurement: Recent Advancements and Noted Limitations. *Mech. Syst. Signal Process.* **2017**, *86*, 35–48. [[CrossRef](#)]
52. Poozesh, P.; Sarrafi, A.; Mao, Z.; Avitabile, P.; Niezrecki, C. Feasibility of Extracting Operating Shapes Using Phase-Based Motion Magnification Technique and Stereo-Photogrammetry. *J. Sound Vib.* **2017**, *407*, 350–366. [[CrossRef](#)]
53. Tian, L.; Zhao, J.; Pan, B.; Wang, Z. Full-Field Bridge Deflection Monitoring with Off-Axis Digital Image Correlation. *Sensors* **2021**, *21*, 5058. [[CrossRef](#)] [[PubMed](#)]
54. Al-Baradoni, N.; Groche, P. Sensor Integrated Load-Bearing Structures: Measuring Axis Extension with DIC-Based Transducers. *Sensors* **2021**, *21*, 4104. [[CrossRef](#)] [[PubMed](#)]
55. Liu, G.; Li, M.; Zhang, W.; Gu, J. Subpixel Matching Using Double-Precision Gradient-Based Method for Digital Image Correlation. *Sensors* **2021**, *21*, 3140. [[CrossRef](#)]
56. Dandois, F.; Taylan, O.; Bellemans, J.; D’hooge, J.; Vandenneucker, H.; Slane, L.; Scheys, L. Validated Ultrasound Speckle Tracking Method for Measuring Strains of Knee Collateral Ligaments In-Situ during Varus/Valgus Loading. *Sensors* **2021**, *21*, 1895. [[CrossRef](#)]
57. De Domenico, D.; Quattrocchi, A.; Alizzio, D.; Montanini, R.; Urso, S.; Ricciardi, G.; Recupero, A. Experimental Characterization of the FRCM-Concrete Interface Bond Behavior Assisted by Digital Image Correlation. *Sensors* **2021**, *21*, 1154.
58. Li, D.S.; Li, H.N.; Fritzen, C.P. The Connection between Effective Independence and Modal Kinetic Energy Methods for Sensor Placement. *J. Sound Vib.* **2007**, *305*, 945–955. [[CrossRef](#)]
59. Li, D.S.; Li, H.N.; Fritzen, C.P. Load Dependent Sensor Placement Method: Theory and Experimental Validation. *Mech. Syst. Signal Process.* **2012**, *31*, 217–227. [[CrossRef](#)]

-
60. Wu, Y.C.; Wu, X.C.; Lu, S.X.; Zhang, J.Q.; Cen, K.F. Novel Methods for Flame Pulsation Frequency Measurement with Image Analysis. *Fire Technol.* **2012**, *48*, 389–403. [[CrossRef](#)]
 61. Rubner, Y.; Tomasi, C.; Guibas, L.J. The Earth Mover's Distance as a Metric for Image Retrieval. *Int. J. Comput. Vis.* **2000**, *40*, 99–121. [[CrossRef](#)]