Event-based Vibration Measurement via Dynamic Mode Decomposition

Hyuga Okamoto*, Kosuke Kurihara*, Yoshihiro Maeda†, Takayuki Hamamoto*
*Tokyo University of Science, Tokyo, Japan, †Shibaura Institute of Technology, Tokyo, Japan
4325522@ed.tus.ac.jp, kuriharak@rs.tus.ac.jp, ymaeda@shibaura-it.ac.jp, hamamoto@rs.tus.ac.jp

***Abstract*—Event cameras can capture scene dynamics at the microsecond level, enabling vibration measurement that contains a high-frequency component. In this paper, we propose an event vision-based vibration measurement method via dynamic mode decomposition (DMD), a spatio-temporal analysis method for multidimensional time-series data based on the linear dynamical model. By analyzing event data using DMD, we can extract vibration components with the aid of physics models.**

***Index Terms*—Dynamic mode decomposition, Event camera, Non-contact vibration measurement.**

I. Introduction

Vibration measurement is fundamental for structural health monitoring [1] and mechanical fault diagnosis [2]. To measure vibration, traditionally, contact-type sensors such as accelerometers and piezoelectric sensors have been used [3]–[5]. These contact-type sensors are often unsuitable for inaccessible targets, such as high-voltage or high-temperature equipment, and culturally significant artifacts. Moreover, physical contact between the sensor and the vibrating object alters vibration characteristics due to mass-loading effects [6], [7], thereby degrading measurement accuracy.

Non-contact vision-based vibration measurement methods have attracted considerable attention [8]–[19]. These methods estimate vibration by analyzing temporal variations in the light reflected from an object, allowing measurement without physical contact. They not only address the limitations of contact-based sensors, but also facilitate full-field vibration analysis by leveraging the spatio-temporal information captured by cameras.

Several vision-based vibration measurement methods that analyze high-speed video have been proposed [8]–[10]. Based on high-speed imaging, these methods can capture high-frequency vibration components that are difficult to detect using standard video cameras that typically operate at 30 fps. Davis *et al*. [8] proposed a method that analyzes phase changes in high-speed video. They construct an image pyramid for each frame and then apply spatial filters to analyze spatial phase changes due to vibration. Although this method is effective for high-frequency vibration analysis, its practical applicability is limited due to the high cost of high-speed cameras and the substantial computational resources required to process high-frame-rate video.

To overcome this issue, event vision-based methods have been proposed [11]–[13]. Event cameras asynchronously sample brightness changes in each pixel, enabling rapid and efficient capture of scene dynamics [20]. Dorn *et al*. [13] proposed a phase analysis method using a complex Gabor filter. They analyzed vibration-derived phase changes in the spatial domain for each pixel and extracted vibration components by applying principal component analysis (PCA).

Although event data essentially contain vibration-related spatio-temporal structures, conventional methods extract dominant vibration components by merely aggregating spatial information. Incorporating physics knowledge of the vibration phenomenon into vibration measurement would enable a physically interpretable extraction with high accuracy.

In this paper, we propose an event vision-based vibration measurement method on the basis of dynamic mode decomposition (DMD), a framework for analyzing spatio-temporal structure from multidimensional signals [21]–[25]. The main contribution of this paper is the incorporation of a physics model into event-based vibration measurement through DMD. DMD is a spatio-temporal modal decomposition method that assumes linear dynamical system. Based on the DMD framework, we can extract vibration components through analysis of the dynamical system of the vibration contained in the event data.

II. Proposed Method

Fig. 1 shows an overview of our proposed method. The event data at the pixel position $\boldsymbol{p}$ and the time τ, denoted by $E(\boldsymbol{p}, \tau)$, contains polarity information $\{+1, -1, 0\}$, signifying an increase, decrease, or no change in brightness, respectively. The proposed method extracts vibration components from $E(\boldsymbol{p}, \tau)$ based on the DMD framework.

A. Observation Matrix Construction

We partition the event data into uniform spatial and temporal grids along both temporal and spatial axes, and calculate the total value of events within each spatio-temporal voxel. This aggregation produces an observation matrix that preserves

This research was supported by the JSPS Grants-in-Aid for Scientific Research (JP24K02964, JP24K23903, JP25K21231).

This article has been accepted for publication in IEEE SENSORS 2025. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/SENSORS59705.2025.11330921

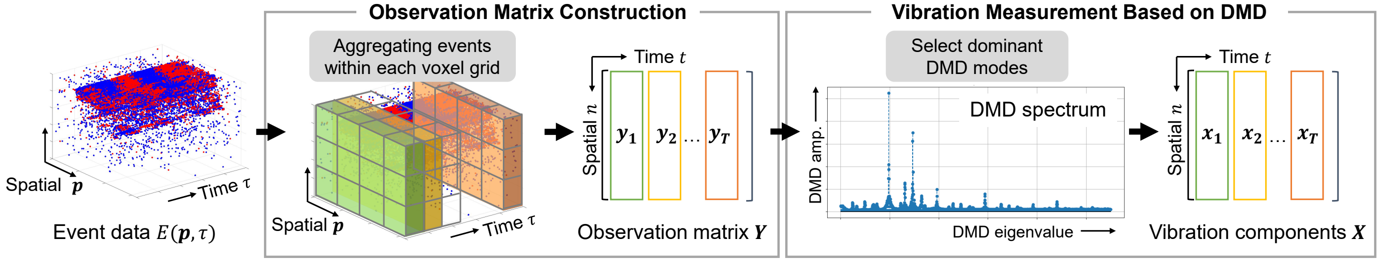


Fig. 1: Overview of our proposed method

amplitude information while mitigating the noise in the raw event data.

Let the indices for the spatial and temporal grids be $n = \{1, \ldots, N\}$ and $t = \{1, \ldots, T\}$, respectively. The (n,t)-th component of the resulting matrix $\boldsymbol{Y}' \in \mathbb{R}^{N \times T}$, denoted by $[\boldsymbol{Y}']_{n,t}$, is obtained as

$$[\boldsymbol{Y}']_{n,t} = \sum_{\boldsymbol{p} \in \mathcal{S}_n} \sum_{\tau \in \mathcal{T}_t} E(\boldsymbol{p}, \tau) \ , \tag{1}$$

where $\mathcal{S}_n$ and $\mathcal{T}_t$ denote the set of spatial and temporal indices corresponding to the n-th spatial grid and the t-th temporal grid, respectively.

By subtracting the temporal average of $\boldsymbol{Y}'$ for each n, we obtain the observation matrix $\boldsymbol{Y} \in \mathbb{R}^{N \times T}$. The t-th column of $\boldsymbol{Y}$, denoted by $\boldsymbol{y}_t \in \mathbb{R}^N$, is obtained as

$$\boldsymbol{y}_t = \boldsymbol{y}'_t - \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{y}'_t \ , \tag{2}$$

where $\boldsymbol{y}'_t$ denotes the t-th column of $\boldsymbol{Y}'$.

B. Spatio-temporal Analysis Based on DMD

Based on the DMD framework, the sequence of $\{\boldsymbol{y}_t\}_{t=1}^{T}$, denoted by $\boldsymbol{Y}_{1:T} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T] \in \mathbb{R}^{N \times T}$ is modeled under the assumption of a linear dynamical system: $\boldsymbol{Y}_{2:T} \approx \boldsymbol{F}\boldsymbol{Y}_{1:T-1}$, where $\boldsymbol{F} \in \mathbb{R}^{N \times N}$ denotes a state transition matrix representing a linear discrete dynamical system. By analyzing $\boldsymbol{F}$, we can decompose $\boldsymbol{Y}$ into the set of DMD modes containing spatio-temporal structures governed by the linear dynamical system.

To obtain $\boldsymbol{F}$, we solve the least squares problem as

$$\boldsymbol{F} = \underset{\boldsymbol{F}}{\arg\min} \, \|\boldsymbol{Y}_{2:T} - \boldsymbol{F}\boldsymbol{Y}_{1:T-1}\|_2^2 \ . \tag{3}$$

We then perform eigendecomposition on $\boldsymbol{F}$ to obtain the J eigenvalues and eigenvectors as $\boldsymbol{F} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^{-1}$, where $\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_J] \in \mathbb{C}^{N \times J}$ is the matrix whose columns are the eigenvectors of $\boldsymbol{F}$. In addition, $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_J) \in \mathbb{C}^{J \times J}$ is the diagonal matrix containing each eigenvalue.

In the DMD framework, the j-th eigenvector $\boldsymbol{w}_j$ and corresponding eigenvalue λ_j represent the spatial structure and temporal oscillation with the growth or decay of the j-th DMD mode, respectively, derived from the linear dynamical system. Additionally, $\beta_j = \mathrm{imag}(f_s \ln(\lambda_j))/2\pi$, where $\mathrm{imag}(\alpha)$ and f_s respectively denote the imaginary part operator and sampling frequency, represents the temporal frequency [Hz] of the j-th DMD mode. In the following, we refer to $\boldsymbol{w}_j$, β_j, and λ_j as the DMD eigenvector, DMD frequency, and DMD eigenvalue, respectively.

By using $\{\boldsymbol{w}_j\}_{j=1}^{J}$ and $\{\lambda_j\}_{j=1}^{J}$, $\boldsymbol{y}_t$ can be decomposed as

$$\boldsymbol{y}_t \approx \boldsymbol{F}^{t-1}\boldsymbol{y}_1 = \boldsymbol{W}\boldsymbol{\Lambda}^{t-1}\boldsymbol{b} = \sum_{j=1}^{J} \lambda_j^{t-1} b_j \boldsymbol{w}_j \ , \tag{4}$$

where $\boldsymbol{b} = [b_1, b_2, \ldots, b_J]^\top = \boldsymbol{W}^{-1}\boldsymbol{y}_1 \in \mathbb{C}^J$ denotes the amplitudes of each DMD mode.

C. Vibration Measurement by DMD Mode Selection

A DMD mode with a higher DMD amplitude b_j represents a more dominant spatio-temporal structure in $\boldsymbol{Y}$, which is expected to represent the vibration component. Based on this, we extract the t-th vibration components $\boldsymbol{x}_t \in \mathbb{R}^N$ by selecting DMD modes corresponding to the top p proportion of DMD amplitudes $\{b_j\}_{j=1}^{J}$ as

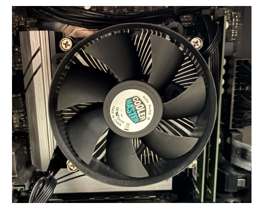
$$\boldsymbol{x}_t = \sum_{j \in \Omega} \lambda_j^{t-1} b_j \boldsymbol{w}_j \ , \tag{5}$$

where Ω is the set of indices of the selected DMD modes.

III. Experiment

To assess the vibration measurement performance, we captured two vibration objects, as shown in Fig. 2. We used an event camera (Prophesee SilkyEvCam VGA) with a resolution of 640×480 and a minimum time resolution of 10^{-6} seconds.

(a) Object #1

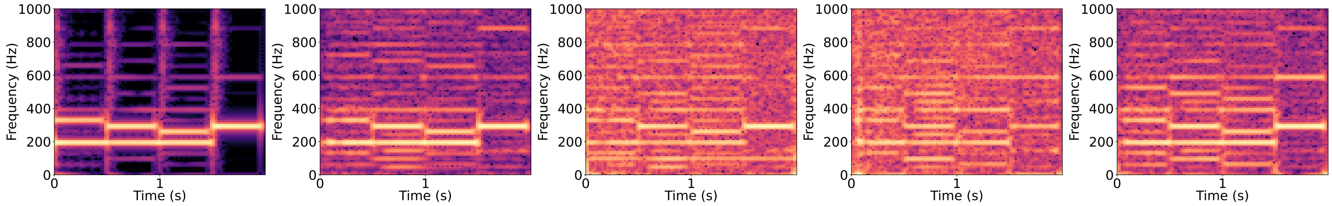


(b) Object #2

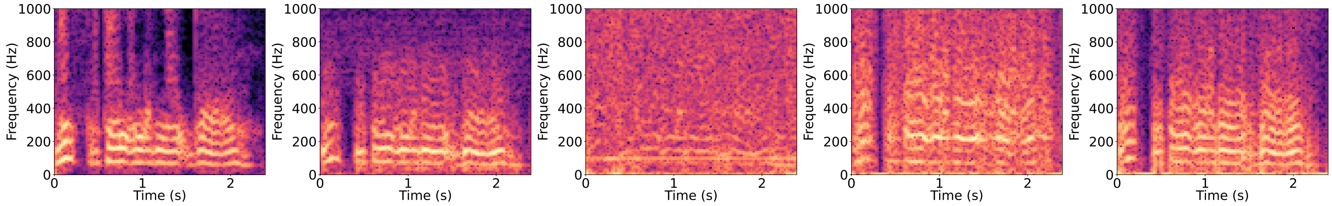
Fig. 2: Measurement objects

A. Experimental Setup

Measurement Object #1: We captured the vibration of an aluminum foil placed on a speaker, as shown in Fig. 2 (a). In this experiment, the foil was vibrated by playing two audio signals provided in a previous study [8]: “Mary” and “Once”.



(a) Ground truth (b) Ours (c) EvSum [11] (d) EvPhase [13] (e) EvPhasePCA [13]



(f) Ground truth (g) Ours (h) EvSum [11] (i) EvPhase [13] (j) EvPhasePCA [13]

Fig. 3: Qualitative evaluation by spectrograms for Object #1. Top row: “Mary”. Bottom row: “Once”.

The audio signals were treated as ground truth vibration signals.

As evaluation metrics, we used the cosine similarity and root mean square error (RMSE) between the amplitude spectrogram of the estimated and the ground truth vibration signals. Since the ground truth is a one-dimensional audio signal while the estimated vibration component $\{\boldsymbol{x}_t\}_{t=1}^{T}$ is multidimensional, PCA was applied to $\{\boldsymbol{x}_t\}$ for dimensionality reduction, treating it as an estimated signal.

For comparison methods, we used three event-based methods: EvSum [11], EvPhase [13], and EvPhasePCA [13]. EvSum reconstructs a vibration signal by simply aggregating event data among all pixels. EvPhase extracts local phase changes derived from vibrations using a complex Gabor filter for a manually selected single pixel. EvPhasePCA extracts the vibration signal by performing PCA on the multidimensional signal obtained by the EvPhase method.

Measurement Object #2: To analyze the capabilities of spatio-temporal analysis of our method, we conducted additional experiments with a simple vibrating object. Specifically, as shown in Fig. 2 (b), we captured a computer fan rotating at a constant speed. This fan has seven blades and rotates at approximately 1600 rpm, resulting in a blade passing frequency of 187 Hz. By analyzing the DMD eigenvectors and eigenvalues, we examined spatio-temporal structures extracted by our method.

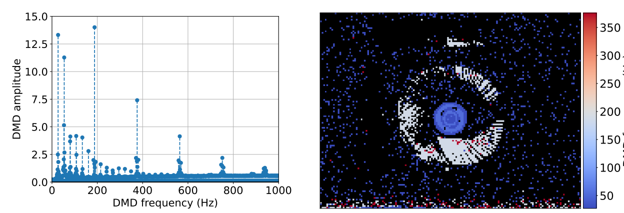
Based on preliminary experiments, we set the number of spatial grids N to 19200. The number of temporal grids T was set such that each temporal interval had a duration of 1/2200 seconds. The proportion for the DMD mode selection p was set to 70 %.

B. Experimental Results

Object #1: Table I shows the comparison results using cosine similarity and RMSE. Our method exhibits the largest cosine similarity and the smallest RMSE for both vibrations, indicating superior performance over other methods. Fig. 3 shows the spectrograms of the extracted vibration signals. Compared to other methods, our method can extract the time-frequency structures of the vibration signals more accurately.

TABLE I: Quantitative comparison for Object #1.

Method	“Mary”		“Once”	
	Cosine sim.	RMSE	Cosine sim.	RMSE
Ours	**0.97**	**0.009**	**0.89**	**0.011**
EvSum [11]	0.75	0.021	0.62	0.018
EvPhase [13]	0.62	0.025	0.59	0.014
EvPhasePCA [13]	0.87	0.014	0.71	0.012



(a) DMD spectrum (b) Top 4 DMD modes

Fig. 4: Experimental results for Object #2.

Object #2: Fig. 4 (a) shows the DMD spectrum, a scatter plot of DMD eigenvalues and the corresponding DMD amplitude. We can see that the DMD frequency with the highest DMD amplitude (187.91 Hz) closely matches the actual blade passing frequency of the computer fan (187 Hz). Fig. 4 (b) visualizes the DMD eigenvectors corresponding to the four highest DMD amplitudes, with each DMD frequency represented by a distinct color. We can see that our method effectively decomposed the spatio-temporal structures of vibration dynamics of the fan such as wings and bearings.

IV. Conclusion

We proposed an event vision-based vibration measurement method based on the DMD framework. By performing DMD on an observation matrix constructed from event data, we can accurately extract spatio-temporal vibration components with a physically meaningful interpretation.

References

[1] Deepam Goyal and BS Pabla, "The vibration monitoring methods and signal processing techniques for structural health monitoring: a review," *Archives of Computational Methods in Engineering*, vol. 23, pp. 585–594, 2016.

[2] Mohamad Hazwan Mohd Ghazali and Wan Rahiman, "Vibration analysis for machine monitoring and diagnosis: A systematic review," *Shock and Vibration*, vol. 2021, no. 1, pp. 1–25, 2021.

[3] Marcus Varanis, Anderson Silva, Arthur Mereles, and Robson Pederiva, "Mems accelerometers for mechanical vibrations analysis: A comprehensive review with applications," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 40, pp. 1–18, 2018.

[4] Alessandro Sabato, Christopher Niezrecki, and Giancarlo Fortino, "Wireless mems-based accelerometer sensor boards for structural vibration monitoring: A review," *IEEE Sensors Journal*, vol. 17, no. 2, pp. 226–235, 2017.

[5] Tianqiong Wu, Di You, Huayun Gao, Pinghua Lian, Weigang Ma, Xinyi Zhou, Chuanmin Wang, Jianghai Luo, Haibo Zhang, and Hua Tan, "Research status and development trend of piezoelectric accelerometer," *Crystals*, vol. 13, no. 9, 2023.

[6] Orhan Cakar and Kenan Y. Sanliturk, "Elimination of transducer mass loading effects from frequency response functions," *Mechanical Systems and Signal Processing*, vol. 19, no. 1, pp. 87–104, 2005.

[7] Orkun Özşahin, H. Nevzat Özgüven, and Erhan Budak, "Analysis and compensation of mass loading effect of accelerometers on tool point frf measurements for chatter stability predictions," *International Journal of Machine Tools and Manufacture*, vol. 50, no. 6, pp. 585–589, 2010.

[8] Abe Davis, Michael Rubinstein, Neal Wadhwa, Gautham J. Mysore, Frédo Durand, and William T. Freeman, "The visual microphone: passive recovery of sound from video," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, pp. 1–10, 2014.

[9] Jaka Javh, Janko Slavič, and Miha Boltežar, "The subpixel resolution of optical-flow-based modal analysis," *Mechanical Systems and Signal Processing*, vol. 88, pp. 89–99, 2017.

[10] Boualem Merainani, Bian Xiong, Vincent Baltazart, Michael Döhler, Jean Dumoulin, and Qinghua Zhang, "Subspace-based modal identification and uncertainty quantification from video image flows," *Journal of Sound and Vibration*, vol. 569, pp. 1–17, 2024.

[11] Friedhelm Hamann, Suman Ghosh, Ignacio Juárez Martínez, Tom Hart, Alex Kacelnik, and Guillermo Gallego, "Fourier-based action recognition for wildlife behavior quantification with event cameras," *Advanced Intelligent Systems*, vol. 7, no. 2, pp. 1–14, 2024.

[12] Matthew Howard and Keigo Hirakawa, "Event-based visual microphone," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.

[13] Charles Dorn, Sudeep Dasari, Yongchao Yang, Charles Farrar, Garrett Kenyon, Paul Welch, and David Mascareñas, "Efficient full-field vibration measurements and operational modal analysis using neuromorphic event-based imaging," *Journal of Engineering Mechanics*, vol. 144, no. 7, pp. 0401805401–0401805412, 2018.

[14] Woong jae Na, Kyung Ho Sun, Byeong Chan Jeon, Jaeyun Lee, and Yun ho Shin, "Event-based micro vibration measurement using phase correlation template matching with event filter optimization," *Measurement*, vol. 215, pp. 1–11, 2023.

[15] Mingzhe Zhao, Xiaojun Shen, and Fei Jiang, "Research on mechanical vibration measurement method based on event camera," in *Proc. International Conference on Energy Engineering and Power Systems (EEPS)*, 2023, pp. 528–532.

[16] Ruiyi Guang, Xiang Li, Yaguo Lei, Naipeng Li, and Bin Yang, "Non-contact machine vibration sensing and fault diagnosis method based on event camera," in *Proc. Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, 2023, pp. 1–5.

[17] Xiang Li, Shupeng Yu, Yaguo Lei, Naipeng Li, and Bin Yang, "Intelligent machinery fault diagnosis with event-based camera," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 1, pp. 380–389, 2024.

[18] Chenyang Shi, Ningfang Song, Boyi Wei, Yuzhen Li, Yibo Zhang, Wenzhuo Li, and Jing Jin, "Event-based vibration frequency measurement with laser-assisted illumination based on mixture gaussian distribution," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–13, 2023.

[19] Muhammad Aitsam, Gaurvi Goyal, Chiara Bartolozzi, and Alessandro Di Nuovo, "Vibration vision: Real-time machinery fault diagnosis with event cameras," in *Proc. European Conference on Computer Vision Workshops (ECCVW)*, 2024, pp. 293–306.

[20] Guillermo Gallego, Tobi Delbruck, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J. Davison, Jorg Conradt, Kostas Daniilidis, and Davide Scaramuzza, "Event-based vision: A survey," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 44, no. 1, pp. 154–180, 2022.

[21] Kosuke Kurihara, Yoshihiro Maeda, Daisuke Sugimura, and Takayuki Hamamoto, "Spatio-temporal structure extraction of blood volume pulse using dynamic mode decomposition for heart rate estimation," *IEEE Access*, vol. 11, pp. 59081–59096, 2023.

[22] Bingni W. Brunton, Lise A. Johnson, Jeffrey G. Ojemann, and J. Nathan Kutz, "Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition," *Journal of Neuroscience Methods*, vol. 258, pp. 1–15, 2016.

[23] Peter J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010.

[24] J. Nathan Kutz, Steven L. Brunton, Bingni W. Brunton, and Joshua L. Proctor, *Dynamic mode decomposition: data-driven modeling of complex systems*, SIAM, Philadelphia, PA, USA, 2016.

[25] Peter J. Schmid, "Dynamic mode decomposition and its variants," *Annual Review of Fluid Mechanics*, vol. 54, pp. 225–254, 2022.