Improving Phase-based Motion Magnification with Lagrangian Synthesis

***Abstract*—Subtle motions in video, such as vascular pulses or mechanical vibrations, are often imperceptible yet critical in medical, structural, or scientific applications. While recent learning-based methods have shown impressive results, phase-based Eulerian motion magnification remains widely used for its subpixel precision, explicit frequency control, and interpretability without the need for training. However, traditional phase-based approaches often suffer from artifacts like ringing and blurring due to the global linearity of phase manipulation.**

In this work, we propose a hybrid method that preserves the temporal precision and interpretability of phase-based Eulerian analysis while reducing its typical artifacts through a Lagrangian-inspired warping synthesis. Unlike previous hybrid approaches that rely on optical flow from artifact-prone magnified videos, our method derives motion representations directly from the phase domain, enabling stable and coherent reconstruction. The framework supports real-time processing, spatiotemporal smoothing, and directional control. Qualitative results show that our method matches the magnification quality of traditional phase-based techniques, significantly reduces visual artifacts, and performs competitively with recent learning-based approaches.

***Index Terms*—Motion magnification, Eulerian analysis, Lagrangian synthesis, Phase-based**

I. Introduction

Subtle motions in videos, such as the expansion of arteries from a heartbeat or minute vibrations of mechanical structures, are often imperceptible to the naked eye. Video motion magnification is a technique that makes these small motions visible by amplifying them in a perceptually meaningful way. First introduced by Wu et al. [1] through an Eulerian framework, motion magnification has since been adopted in a wide variety of domains—from medical diagnostics and structural monitoring to human-computer interaction and surveillance []. Its appeal lies in turning standard video footage into a powerful analytical tool, enabling insights that would otherwise require specialized sensing equipment.

Among the different strategies for motion magnification, phase-based methods have proven particularly robust and interpretable. Proposed by Wadhwa et al. [2], the phase-based approach leverages the shift-invariance property of phase in complex steerable pyramids to extract and amplify motion signals with subpixel accuracy. It provides explicit control over spatial and temporal frequency bands, making it ideal for amplifying well-defined periodic motions such as heartbeats or vocal cord vibrations [2].

In recent years, deep learning-based motion magnification methods have emerged, often achieving impressive results through end-to-end learning [3]–[7]. However, these models are frequently trained on specific motion types and may not generalize well across varying frequencies or contexts. Moreover, many do not include an explicit temporal filtering stage, which is critical in traditional hand-crafted Eulerian frameworks for selecting motion of interest. As a result, more classical, phase-aware techniques remain highly relevant in scientific, medical, and engineering domains where temporal precision, interpretability, and frequency isolation are crucial.

Among the Eulerian hand-crafted methods [1], [2], [8]–[10], the phase-based [2] method remain state-of-the-art in high-precision applications. Recent studies demonstrate the effectiveness of the phase-based method in real-world scenarios such as vascular pulse enhancement in robotic surgery [11], vibration analysis of structural elements [12], and non-contact infrastructure monitoring [13], [14] among other applications [15], [16]. They provide superior temporal fidelity, field-wide vibration capture, and interpretable motion representation—without relying on complex training pipelines. Innovations in filter design and phase estimation have further improved accuracy in micro-vibration measurement, structural mode shape extraction, and pulsation magnification. However, due to the use of the phase-based in these method, the results usually present ringing artifacts, blurring. therefore, challenges that learning-based approaches attempt to address, albeit often at the expense of transparency and adaptability.

In this paper, we propose a novel method that maintains the advantages of phase-based Eulerian analysis while mitigating its most common artifacts and advantages of Lagrangian synthesis stage. By combining Eulerian phase analysis with a Lagrangian-inspired warping synthesis, our approach captures subtle phase changes in motion and reconstructs perceptually coherent, high-quality videos. The method eliminates the need for end-to-end learning and instead leverages the robustness of steerable pyramids and phase manipulation to obtain a motion flow and use it to reconstruct a motion magnified video with reduced ringing and blurring.

Our motivation stems from the observation made in previous works [16]–[18], where motion vectors estimated (via optical flow) between an original and a magnified video frames were used to synthesize clean, artifact-free outputs via warping. These works demonstrated how separating motion analysis (e.g., magnification, filtering) from motion synthesis (e.g., warping) allows greater flexibility and avoids entangling visual appearance with motion representation. However, these methods depends critically on a high-quality input magnification; if the initial amplified video contains temporal artifacts or spatial distortions, the resulting motion vectors become unreliable due a poor correspondence of the optical flow. In contrast, the present work adopts a similar strategy on the video reconstruction via warping, but do not rely on the optical flow of synthesized videos with possible artifacts. Instead, we obtain the motion representation directly from the phase space.

Our implementation presents a also adds the analysis on different temporal motion modes (static, dynamic and filtering), spacial smoothing for motion noising handling, orientation control and processes videos in real time. We demonstrate by qualitative video analysis that the proposed method is able to achieve higher magnification factors than learning-based methods (similar to phase-based magnification), while eliminating the visual artifacts.

The main contributions of out method are:

- A new method for video motion magnification based on the analysis of phase space of complex steerable pyramids, and leveraging Lagrangian synthesis and its advantages on post-processing options, artifact-free reconstruction, interpretability and high factors of magnification. Our approach is a hand-crafted method, and it is hybrid, since it is formed by Eulerian Analysis and Lagrangian synthesis.
- A real time implementation of our method, using GPU acceleration, via Pytorch, where all parameters of the method can be modified in real time.
- A set of qualitative experiments in which we compare the proposed method to recent state-of-the-art works, specially learning-based approaches. When compared to these methods, we show that our method is capable for higher magnification factors and can greatly decrease the presence of artifacts, even when compared to recently proposed learning-based methods.

II. Motion Magnification Paradigms

The literature of the video motion magnification usually present some paradigms/philosophies. The main classes of operations are divided into Eulerian vs Lagrangian methods and Hand-crafted vs learning based methods.

A. Lagrangian vs Eulerian

Lagrangian motion magnification methods [5], [19], [20] explicitly track motion trajectories over time by estimating displacement fields through optical flow or sparse feature tracking. These displacements are then used to synthesize magnified motion by warping video frames accordingly. While Lagrangian techniques tend to be more accurate in scenarios involving large or complex motions, they rely heavily on the quality of motion estimation. As a result, their performance can degrade significantly in the presence of noise, low texture, or occlusion, where tracking becomes unreliable.

Eulerian motion magnification techniques amplify subtle temporal changes at fixed spatial locations without tracking motion explicitly. The first proposed Eulerian method in the literature [1] analyses the motion in intensity variations. Next, the Phase-based methods [2], [8], improved robustness by manipulating local phase shifts in multiscale decompositions, enabling much higher motion magnification and noise suppressing. Other approaches include the use of handcrafted temporal filters, such as acceleration [9] and jerk [10] to target specific motion characteristics. The Eulerian approaches usually presents ringing and blurring under strong magnification, as consequence of the manipulation on the Fourier space.

B. Hand-crafted vs Learning-Based Methods

The previously presented methods con be classified as hand-crafted methods, they filters are pre-determined or designed to tailor specific needs.

Otherwise, learning-based methods [4]–[7] use neural networks to automatically learn the filters that best represent the motion magnification phenomenon. Most of them are Eulerian, where there is a encoder and decoder that represent the shape and textures and reconstruct the video with magnification at the end. Although these methods mitigate a lot of the ringing and blurring artifacts, learning based models tend to be data-dependent, less interpretable, and may generalize poorly to unseen motions or frequencies. They typically lack the precise temporal control offered by phase-based Eulerian pipelines and may hallucinate artifacts or amplify undesired motions.

This motivates a return to physically interpretable frameworks, where temporal frequency and phase control are directly handled. Our method inherits this transparency, offering analytical interpretability with improved synthesis quality. Also, there is a trend on the literature, where many recent works are preferring to use hand-crafted, Eulerian approaches, such Phase-based [11]–[16].

III. Phase-Based Analysis and Lagrangian Synthesis

Our method is a hand-crafted method, based on Eulerian phase-based motion analysis for robust and Lagrangian based video synthesis for a robust motion magnification. We call it Phase-Based with Lagrangian synthesis, or **PBLS** for short. Figure 1 presents a simplified representation of our method. In next sections we describe the method in details.

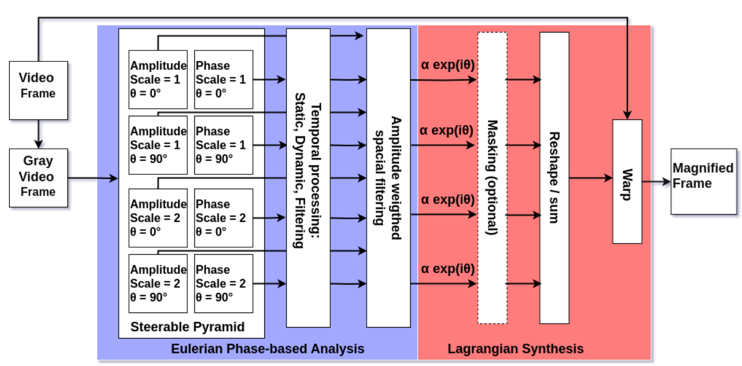

Fig. 1. The diagram of our proposed Phase-based with Lagrangian Synthesis method for motion magnification. It is divided in a analysis stage, which is very similar to the original phase-based method, and a Lagrangian synthesis.

A. Motivation

This works is greatly inspired by the Synflowmap method [17], [18]. It uses a magnified video with an Eulerian method and the original video to obtain the motion estimation of the magnified video compared to the original (analysis stage), and applies the warp to the original video based on the estimated motion (synthesis stage).

In this work, we borrows these ideas and apply directly in the intermediate stages of the Phase-based method. In phase-based method, you obtain an internal representation that contains the magnified motion of the video (in terms of a pyramidal steerable transform). In this work, we translate and combine the coefficients directly to motion information, and apply the motion directly to the video input, like in the synthesis of Lagrangian methods.

In learning based methods, it is very common to represent the input videos in terms of the texture and shape representations, in which the video motion is encoded on the shape representation while the intensity is in the texture representation. The learning based method rely on the correct reconstruction of both representations. We think the texture representation could be extracted directly from the original video, with no need of any reconstruction. It maintains exactly the same aspect of the original video, elimination every reconstruction artifact. If the motion is correctly extracted, our hypothesis is that this approach is the recipe to a artifact-free motion magnification.

B. Problem formulation

Lets first describe the problem. We use the same motion magnification formulation used in many state of the art-works [1], [2], [4], [18].

Lets say we have a video

$$I_o(\mathbf{x},t) = f(\mathbf{x} + \delta(\mathbf{x},t)), \tag{1}$$

where $\delta(\mathbf{x},t)$ is the motion function for each position $\mathbf{x}$ and time, and $f(.)$ is the function that represents the video pixels.

Eulerian motion magnification aims to synthesize a video that amplifies the displacement T($\delta(\mathbf{x},t)$) by a factor α, where $T(.)$ selects motion of interest. The video with magnified motion is represented by:

$$I_m(\mathbf{x},t) = f\left(\mathbf{x} + \delta(\mathbf{x},t) + \alpha T(\delta(\mathbf{x},t))\right). \tag{2}$$

1) Eulerian Analysis: The analysis stage consists of obtaining the magnified representation of the video signal. The goal is to find an approximate motion of the current video at the time stamp t, and multiply by a factor of α, as given by the following equation:

$$\hat{\delta}(\mathbf{x},t) \approx \alpha T(\delta(\mathbf{x},t)).$$

In our work, we use the complex steerable pyramid decomposition [21], [22], similarly to the Phase-based and acceleration magnification methods [2], [9]. It is an over-complete representation that contains the amplitude and phase of Fourier coefficients in many scales and orientations. It was very well documented by previous methods that the phases carries the displacements, and by applying an operation in time, it carries a representation of the motion, that can be magnified.

Let $\phi_{s,\theta}(x,t)$ the phase of the complex steerable pyramidal decomposition of the video, at the scale $s \in \{2^0, 2^1, ..., 2^{S-1}\}$, orientation $\theta \in \{\theta_0, \theta_0 + \frac{\pi}{N_\theta}, \theta_0 + \frac{2\pi}{N_\theta}, ..., \theta_0 + \frac{(N_\theta - 1)\pi}{N_\theta}\}$, and position x and time t. Our work rely on the assumption that the desired motion flow $F_{s,\theta}$ is similar to the phase after applying the motion selection scheme:

$$F_{s,\theta}(x,t) \approx T(\phi_{s,\theta}(x,t)$$

Notice that there is a motion flow for each scale and orientation.

The original learning-based work [4] defined the static, dynamic and filtering modes of motion $T(.)$, where the static correspond to the difference of motion between the current frame I_{t_k} a anchor frame (usually the first I_0), dynamic to the difference to the previous frame $I_{t_{k-1}}$, and filtering is a mode of motion defined by a temporal filter of previous frames. The temporal filtering allows to select the frequency band of the motion, to better tailor to specific uses of motion magnification, like isolating a specific motion. In our method, follows this formulation, allowing for the 3 types of motion modes. We define the motion selection as follows:

$$\begin{aligned} \text{Static mode:} \quad & F_{s,\theta}(x,t_k) = \phi_{s,\theta}(x,t_k) - \phi_{s,\theta}(x,t_0) \\ \text{Dynamic mode:} \quad & F_{s,\theta}(x,t_k) = \phi_{s,\theta}(x,t_k) - \phi_{s,\theta}(x,t_{k-1}) \\ \text{Filter mode:} \quad & F_{s,\theta}(x,t_k) = \mathcal{H}\phi_{s,\theta}(x,t_k). \end{aligned} \tag{3}$$

Here, the time stamp t_k refers to the time at the frame k. $\mathcal{H}$ represent a general temporal filter. As F manipulates the phase signal, a signal that is expected to be limited to $-\pi$ and π, we wrap the angle back to this interval after applying F.

Naturally, signal is expected to present noise, which can lead to incorrect motions being amplified. To handle with noise, we employed similar strategy to phase-based method [2]. We apply a spatial low-pass Gaussian filter to both the F output and the amplitude signals. The phase is filtered with amplitude-based weighting, since the phase in low-amplitude regions is usually not meaningful. The final output is computed as the ratio between the filtered phase and the filtered amplitude. To control the level of spatial smoothing, the standard deviation σ_1 determines the extent of the smoothing and is treated as a parameter of our method.

Formally, we can represent this operation by

$$\hat{F}_{s,\theta}(x,t_k) = \frac{\mathcal{G}_{\sigma_1}[F_{s,\theta}(x,t_k)]}{\mathcal{G}_{\sigma_1}[A_{s,\theta}(x,t_k)]},$$

where $\mathcal{G}_\sigma$ is the Gaussian smoothing.

C. Lagrangian Synthesis

After obtaining the motion flow $\hat{F}_{s,\theta}$ we can process, magnify and synthesize the video. Because it is a Lagrangian synthesis method, it allows to apply processing techniques before reconstructing the video, instead of directly apply the pyramidal reconstruction.

We apply another filtering step on the isolated motion $\hat{F}_{s,\theta}$. We tested both a Gaussian low pass and a bilateral filtering, to the user to choose which to use. The standard deviation of the second smoothing is represented by σ_2.

Because it is a Lagrangian approach, our method allows us to apply a mask where we want to apply the magnification, similarly of the proposed on flowmag method [5]. Therefore, if the user want to magnify only on a specific object, the user provides the mask that selects this object. If doing so, we apply a morphological opening to the input mask and multiply $\hat{F}_{s,\theta}$ by 0 in the region outside the opened mask. We observed that the opening operation gives a better margin and avoids discontinuities.

The magnification happens next. The signal $\hat{F}_{s,\theta}$ contains the estimated motion for each scale and angle. so, we magnify its amplitude and translate it to the correct angle and scale. The scales are always by a factor of 2^s, so, at a scale s we resize the motion flow by a factor of 2^s. At every orientation θ ($\theta \in \{\theta_0, \theta_0 + \frac{\pi}{N_\theta}, \theta_0 + \frac{2\pi}{N_\theta}, ..., \theta_0 + \frac{(N_\theta - 1)\pi}{N_\theta}\}$), we rotate the delta by multiplying it by $e^{i\theta}$.

The total combined motion flow can be represented by:

$$F(x,t) = \alpha \sum_s \sum_\theta \hat{F}_{s,\theta}\left(\frac{-\mathbf{x}}{2^s}, t\right) \cdot e^{i\theta} \tag{4}$$

Notice a negative factor multiplying x in this equation, this correspond to the 180 degree rotation to compensate by the decomposition.

Finally, we warp the original frames of the video by the calculated motion flow.

Let $W(I,\delta)$ be the warp function that maps the image $I = f(\mathbf{x})$ to the image $f(\mathbf{x}+\delta)$. Our final output:

$$I_m(\mathbf{x},t) = W(I(\mathbf{x},t), F(\mathbf{x},t)) \tag{5}$$

Figure 2 shows the motion vectors as $\alpha \hat{F}_{s,\theta}\left(\frac{-\mathbf{x}}{2^s}, t\right) \cdot e^{i\theta}$ drown above the a given frame image of the original image. The Figure represents all 4 scales and 2 orientations. The first orientation is 0 and the second is $\pi/2$. Notice that the firs column of images has only arrows in the horizontal direction and the second only vertical ones, according to the selected orientation. Notice also that most of the magnification is perceived images with higher downscale. The final synthesized frame is obtained by getting these motion representation scaling and summing, and use the result to warp the original frame. We observed that indeed, this synthesized video present very similar motion amplification to the phase-magnification pyramidal decomposition reconstruction.

D. Implementation details

This section present more details of the implementation proposed in this work.

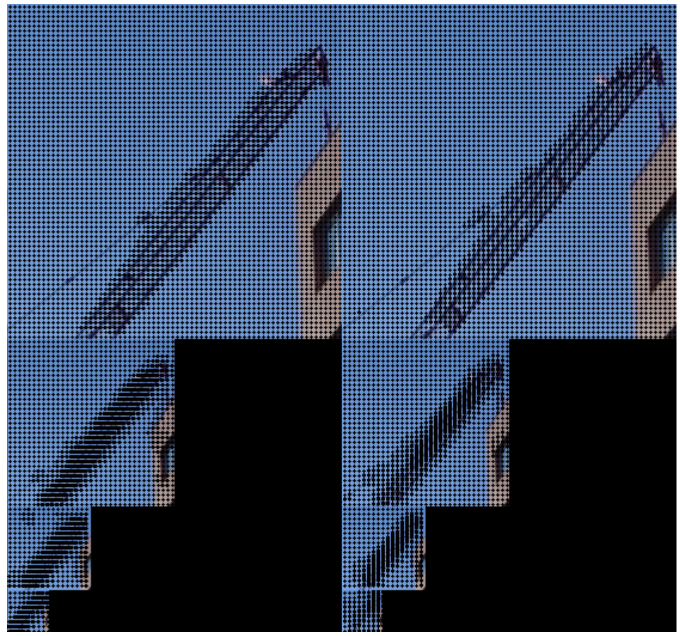

Fig. 2. Motion vectors for each scale and orientation

1) Real time processing: We developed the method using a Python implementation, and using Pytorch to take advantage of parallel processing of a GPU. We prepared our implementation for real time usage, each frame is processed alone at a time. We only store other frames decomposition for calculating the static/dynamic differences. In the case of filtering mode, we used a IIR filter and keep track of necessary delayed inputs/outputs. In fact, the filter can be changed in real time. We prepared it to the user select the passing band in real time.

2) Spacial filtering: The default σ_1 and σ_2 values for the Gaussian smoothing is 5. If they are too big, the amplification is softened and translated to surrounding areas, but if too small, the noise is usually not suppressed. We observed that the value of 5 represent a good compromise. This value can also be changed by the user in real time.

3) Pyramidal decomposition: The complex steerable pyramidal decomposition is applied at every frame, using a Pytorch implementation [1] of the complex steerable pyramid. Regarding on the parameters of the decomposition, we observed that 2 orientations is enough to represent the motion in any direction, while phase-based usually require 8. Also, the method allows to user to isolate the motion on a single orientation and reconstruct the magnified video based on this orientation. The number of orientations and angle of a single orientations can be set in real time as well. Regarding on the decomposition levels, we set between 3 and the maximum possible to the image resolution.

While the phase-based method usually uses half or quarter-octave decomposition to decrease the strength of ringing and blurring like artifacts, our method can effectively get rid of these artifacts easily. Therefore, octave level decomposition

[1] https://github.com/tomrunia/PyTorchSteerablePyramid

is enough for our method, what decreases the computational burden and allows for faster processing.

IV. Comparison to state-of-the-art

A. Conceptually

Our work can be view by multiple viewpoints. Firstly, when compared to Phase-based [2], it substitutes the pyramidal decomposition by viewing the processed phase space as a motion field, and directly applies warp to the original video based on it. It re-create the composition as a combination scales and warp operations from estimated motion flow and the original video. Our method also do not need the low or high frequency residuals, and also do not need the reconstruction of the YIQ color space. We can evaluate the motion field in on a single gray scale channel and translate the motion to original colors in the original video. Our method also includes static and dynamic temporal processing in the phase-based analysis.

When compared to the Synflowmap [18], instead of obtaining the motion flow from a pre-calculated Eulerian magnification method and the original video frames and applying it on the original video, our method calculates the motion flow before synthesizing an Eulerian magnified video. Therefore, despite the similarity between the synthesis stage of both methods, our method do not rely on an off-the-shelf pre-calculated magnified video.

Eulerian Learning-based methods such the deepmag or EulerMormer [4], [7] usually present a latent representations of texture and shape using learned encoders. The neural networks have to reconstruct both the motion information and the texture representation. That requires the models to fairly represent the structures of the natural world that appear on the videos. Our method aims to represent the motion information, given all texture information is already present on the original video.

B. Qualitative results

We evaluate the quality of the video motion magnification in a set of videos, and we compare to the results of methods **Phase-based** [2], the original learning-based method, named **deepmag** [4], and two recent approaches, the **flowmag** [5] and the **EulerMormer** [7]. Table I summarizes the features of these methods.

The Phase-based results were obtained as a composition of the pyramidal representation obtained from our method. Therefore, it uses the same pyramidal parameters: 2 orientations and octave pyramids, the spacial Gaussian filtering, and the same temporal selection mode as used in our method. The deepmag, flowmag and EulerMormer were obtained from the implementation provided by the authors, with default settings. These methods have both static and dynamic modes, but only deepmag has a filtering mode.

We used a dozen videos that are commonly used for qualitative evaluation of motion magnification methods, either small motion (static video dataset)[2] and videos with presence of large motion (dynamic video dataset) [3].

Our approach is qualitative, as this dataset do not have the ground truth videos containing magnified motion. In fact, on the literature, quantitative evaluation with metrics of quality is only used on synthetic videos datasets, usually the ones used for training the learning-based methods.

Figure 3 shows a crop of a frame of the results of the *baby* video sequence, for a magnification factor of 20 and using the static mode. The video contains a baby sleeping, with very slim motions on his belly due to his breath. We can see that all methods can magnify the belly motion, only flowmag shows a smaller magnification. All learning-based methods get rid of the ringing like artifacts compared to phase-based, but each one present different loss of details. Deepmag shows overall blurring on some details, flowmag present noise. Eulermormer mostly preserves well the details of the image, but here we selected a frame in which it fails in reconstruct the eye (it is hardly seen in the video). However, our method seems to better preserve the details of the original image.

Figure 3 shows a crop of two frames of the results of the *cattoy* video sequence, for a magnification factor of 10 and using the dynamic mode. The video contains a mouse toy vibrating and squashing while goes left. Analyzing the image quality, we observe much of ringing and blurring on phase-based method, levels of different levels of blur for the learning methods, while ours maintain the details.

Analyzing the motion, the desired result for the first shown frame would be presenting a contracting to the object, revealing the background, while showing a stretch on the second. The results reveal that it is hard to understand the motion on phase-based result, flowmag didn't get a magnification as the other methods, and deepmag seems to have a little time desynchrony compared to the others. Interestingly, EulerMormer presented the correct expected result in terms of motion, while trying to reconstruct the background (despite very blurred results), also, the eyes seems to pop out of the toy in the second frame, what is not what we would expect. Our method, however, keeps correct track of the motion, but do not handle with disocclusions. The solution on out method is to warp the background to where we would expect to reconstruct the background.

In these results, the magnification factor is relatively small. When we select higher α, the learning based methods start to present stronger visual artifacts. However, our method tends to translate the motion to other objects or background, resulting in incorrect motion model for these objects. However, due to Lagrangian synthesis control, we can use a mask to eliminate background of objects not of interest, and circumvent this issue. The only method that also has similar feature is the flowmag. Also, our method can handle generic filter mode for motion, and it is able to isolate specific motion for specific applications.

[2]https://drive.google.com/drive/folders/1Bm3ItPLhRxRYp-dQ1vZLCYNPajKqxZ1a

[3]https://drive.google.com/drive/folders/1t5u8Utvmu6gnxs90NLUIfmIX0_5D3WtK

Method	class	class	mag. level	motion model	masking
Phase-Based	Eulerian	Hand-crafted	High	filter	no
deepmag	Eulerian	Learnign-based	Medium	static, dynamic, filter	no
flowmag	Lagrangian	Learnign-based (SS)	Low	static, dynamic	yes
EulerMormer	Eulerian	Learnign-based	Medium	static, dynamic and filter	no
PBLS (ours)	Hybrid	Hand-crafted	High	static, dynamic, filter	yes

TABLE I
CAPTION

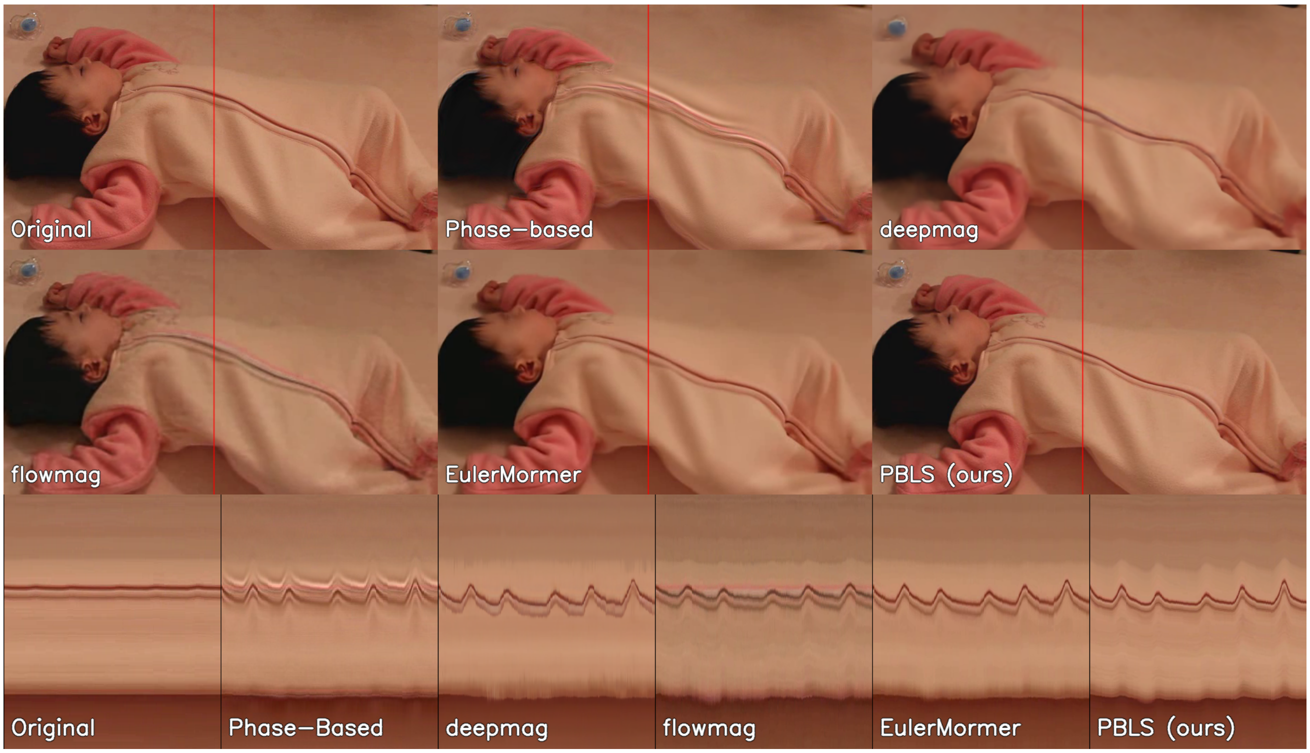

Fig. 3. A frame of the *baby* video sequence. Top 2 rows show the results of motion magnification for evaluated methods. Last row shows the slice-time corresponding to the red line on images above, where the horizontal axis represent time.

Fig. 4. Frame 80 (left) and 85 (right) of the *cattoy* video sequence.

V. Conclusion

We have presented a hybrid motion magnification method that combines the temporal precision and interpretability of phase-based Eulerian analysis with a Lagrangian-inspired warping synthesis. By avoiding reliance on optical flow from artifact-prone magnified videos and instead deriving motion representations directly from the phase domain, our approach achieves high-quality magnification with significantly reduced visual artifacts such as ringing and blurring. Additionally, the method supports real-time processing, temporal selection, orientation-aware control, object selection control, making it both practical and adaptable for real-world applications.

Among the tested methods, experimental results demonstrate that our method produces the best image quality and motion magnification due to original phase-based analysis and direct correspondence to the original content. However, the method usually struggles in occlusions, disoclusions and borders. Overall, it retains the strengths of traditional techniques while performing competitively with recent learning-based approaches, offering a robust solution for scenarios where accuracy, clarity, and interpretability are essential.

REFERENCES

[1] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman, "Eulerian video magnification for revealing subtle changes in the world," in *ACM TOG*, 2012.

[2] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, "Phase-based video motion processing," in *ACM TOG*, 2013.

[3] S. Zhou, S. Hadfield, and R. Bowden, "Deepmag: Source specific motion magnification using gradient ascent," in *ECCV*, 2020.

[4] T.-H. Oh, J.-Y. Lee, N. Xu, C. Kim, M. F. Cohen, and W. T. Freeman, "Learning-based video motion magnification," in *ECCV*, 2018.

[5] Z. Pan, D. Geng, and A. Owens, "Self-supervised motion magnification by backpropagating through optical flow," in *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. [Online]. Available: https://arxiv.org/abs/2311.17056

[6] J. Singh, S. Murala, and G. Kosuru, "Lightweight network for video motion magnification," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 2041–2050.

[7] F. Wang, D. Guo, K. Li, and M. Wang, "Eulermormer: Robust eulerian motion magnification via dynamic filtering within transformer," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 6, 2024, pp. 5345–5353.

[8] D. Sanchez and P. Monasse, "Motion magnification using the riesz pyramid," in *ICIP*, 2018.

[9] Y. Zhang, N. Wadhwa, K. Sunkavalli, M. Chandraker, R. Ramamoorthi, and P. Sen, "Video acceleration magnification," in *ACM TOG*, 2017.

[10] H. Liu, S. Lu, T. Dai, and X. Xie, "Jerk-aware video acceleration magnification," in *ACM MM*, 2020.

[11] Y. Yang and Q. Jiang, "Magnification and localization of vessels in robotic surgical videos based on accuracy high-order phase-based video magnification," *Biomedical Signal Processing and Control*, vol. 96, p. 106575, 2024.

[12] K. Luo, X. Kong, J. Li, J. Hu, and L. Deng, "Motion magnification for video-based vibration measurement of civil structures: A review," *Mechanical Systems and Signal Processing*, vol. 220, p. 111681, 2024.

[13] Y. Yang and Q. Jiang, "A novel phase-based video motion magnification method for non-contact measurement of micro-amplitude vibration," *Mechanical Systems and Signal Processing*, vol. 215, p. 111429, 2024.

[14] N. Guo, S. Wang, J. Lv, S. Cao, and C. Xu, "Damage detection for highly flexible beam-like structures using attenuation response from vision-based vibration measurement," *Available at SSRN 5179258*.

[15] C. Taslicay, S. Occhipinti, J. Herbert, V. Clerc, C. T. d. Cabo, Z. Mao, C. Schwingshackl, and M. R. Brake, "Visualization and identification of nonlinear structural dynamics via phase-based motion magnification," *Available at SSRN 5165980*.

[16] J. Wu, B. Yan, Y. Xue, J. Qin, D. You, and G. Sun, "Multi-point optical flow cable force measurement method based on euler motion magnification," *Buildings*, vol. 15, no. 3, p. 311, 2025.

[17] J. A. Lima, C. J. Miosso, and M. C. Farias, "Hybrid motion magnification based on same-frame optical flow computations," in *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2020, pp. 1–7.

[18] ——, "Synflowmap: A synchronized optical flow remapping for video motion magnification," *Signal Processing: Image Communication*, vol. 130, p. 117203, 2025.

[19] C. Liu, A. Torralba, W. T. Freeman, F. Durand, and E. H. Adelson, "Motion magnification," *ACM transactions on graphics (TOG)*, vol. 24, no. 3, pp. 519–526, 2005.

[20] S. Park and D. Kim, "Subtle facial expression recognition using motion magnification," *Pattern Recognition Letters*, vol. 30, no. 7, pp. 708–716, 2009.

[21] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE transactions on Information Theory*, vol. 38, no. 2, pp. 587–607, 1992.

[22] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International journal of computer vision*, vol. 40, pp. 49–70, 2000.