Eulerian Video Magnification for Revealing Subtle Temporal Variations

Negin Heidarifard

Master Artificial Intelligence, Université Paris-Saclay

Abstract

This project investigates the application of Eulerian Video Magnification (EVM) to enhance subtle temporal variations in facial videos, with a particular focus on physiological signals such as cardiac pulse. By combining color space transformation, multiscale spatial decomposition, temporal bandpass filtering, and region-of-interest masking, the proposed pipeline aims to amplify weak signals while preserving visual realism. The experimental results show that imperceptible color fluctuations in facial regions can be revealed with stable temporal behavior and limited visual artifacts. While effective under controlled conditions, the method also exhibited a notable sensitivity to parameter choices and spatial scaling, which motivated a closer analysis of stability and artifact formation.

1   Introduction

Eulerian Video Magnification (EVM) is a technique designed to amplify small temporal changes in video sequences without explicitly tracking motion. Unlike Lagrangian approaches, which follow individual points over time, EVM operates at fixed spatial locations, making it particularly well-suited for detecting low-amplitude phenomena such as physiological color variations caused by blood flow.

Such subtle signals are difficult to observe with the naked eye, yet they play an important role in applications including non-contact health monitoring and biomedical signal analysis. Beyond reproducing the method, the goal of this project was to understand how EVM behaves in practice, especially with respect to robustness, parameter sensitivity, and visual artifacts that arise outside ideal settings.

2   Methodology

2.1   Pipeline Overview

The complete EVM pipeline consists of the following steps:

1. Loading the input video and extracting frames along with the frame rate.

2. Converting frames from RGB to YIQ color space to separate luminance from chrominance components.

3. Detecting the facial region in the first frame and generating a binary mask, which is kept fixed throughout the sequence.

4. Decomposing each frame into multiple spatial frequency bands using a Laplacian pyramid.

5. Resizing and applying the face mask at each pyramid level.

6. Applying a temporal Butterworth bandpass filter to isolate the target frequency range.

7. Amplifying the filtered signal using a scale-dependent amplification factor.

8. Reconstructing the frames by collapsing the Laplacian pyramid.

9. Converting the processed frames back to RGB and saving the output video.

2.2 Eulerian Magnification Model

At the core of EVM is the idea of amplifying temporal intensity variations at fixed spatial locations. Conceptually, the magnified signal can be expressed as:

$$I'(x,t) = I(x,t) + \alpha \cdot \text{filtered}(I(x,t)), \tag{1}$$

where $I(x,t)$ denotes the original pixel intensity at spatial location x and time t, filtered$(\cdot)$ represents a temporal bandpass-filtered signal, and α is an amplification factor controlling the strength of magnification.

2.3 Scale-Dependent Amplification

To reduce noise amplification at fine spatial scales, the amplification factor is adapted across pyramid levels:

$$\alpha_{\text{level}} = \frac{\alpha}{2^{(\text{pyramid\_levels}-\text{level}-1)}}. \tag{2}$$

This strategy favors amplification of coarser spatial structures while limiting the influence of high-frequency noise, which proved important for visual stability in practice.

3 Results and Analysis

3.1 Amplification of Subtle Color Variations

The magnified videos reveal periodic color fluctuations in facial regions such as the forehead and cheeks. These variations are consistent with blood flow–related pulsations, indicating that the selected frequency band (0.8–1.0 Hz) is well aligned with typical heart rate dynamics. Under moderate amplification, the results remain visually natural, without noticeable flickering or color distortion.

3.2 Temporal Stability

The amplified signal evolves smoothly over time, with no abrupt intensity changes or temporal jitter. This suggests that the Butterworth bandpass filter effectively isolates the desired frequency range, although small parameter changes were observed to significantly affect stability.

3.3 Spatial Artifacts

Minor visual artifacts appear near sharp edges, such as glasses or background boundaries. These effects are mainly attributable to high-frequency pyramid levels capturing fine spatial details that are not directly related to physiological signals. Additionally, thin black borders are visible around some frames, reflecting slight spatial inconsistencies during pyramid reconstruction.

4 Discussion

A key observation from these experiments is the sensitivity of EVM to parameter choices. While the combination of color space separation, multiscale decomposition, and spatial masking improves signal clarity, seemingly reasonable adjustments can quickly introduce artifacts or suppress meaningful variations. This highlights the fragility of the method when applied beyond controlled conditions and underscores the importance of careful tuning and qualitative inspection.

5 Conclusion

This project demonstrates that Eulerian Video Magnification can enhance weak physiological signals in facial videos using a relatively simple pipeline, provided that parameters are carefully selected. More importantly, the experiments reveal that the method is highly sensitive to design choices, motivating a broader interest in robustness and reliability for vision-based systems. These observations connect naturally to questions of decision-making and uncertainty in real-world perception tasks.

References

[1] Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., & Freeman, W. (2012). *Eulerian Video Magnification for Revealing Subtle Changes in the World.* ACM Transactions on Graphics.

[2] Poh, M.-Z., McDuff, D. J., & Picard, R. W. (2010). *Non-contact, Automated Cardiac Pulse Measurements Using Video Imaging.* Optics Express.