Eulerian Video Magnification for Revealing Subtle Changes in Videos

Negin Heidarifard

M2 in Artificial Intelligence, Paris-Saclay University

Abstract

This project explores Eulerian Video Magnification (EVM) through two implementations aimed at unveiling subtle color and motion variations in video sequences. The basic implementation uses RGB-domain processing with simple spatial and temporal filtering. The advanced implementation enhances this by incorporating:

- Color space conversion (RGB $\leftrightarrow$ YIQ) for better luminance preservation,
- Dynamic face detection and masking using Haar Cascades for focused region-of-interest (ROI) amplification,
- Multiscale spatial decomposition using a Laplacian pyramid,
- Adaptive temporal filtering via a zero-phase Butterworth bandpass filter.

The report includes detailed mathematical derivations, algorithmic justifications, and experimental evaluations. Results show that the advanced pipeline achieves clearer visualization of physiological signals (e.g., pulse) and subtle mechanical vibrations while balancing noise reduction and amplification effectiveness.

1. Introduction

Eulerian Video Magnification (EVM) amplifies low-amplitude spatiotemporal signals, such as heartbeats or structural vibrations, without explicit motion tracking. Such imperceptible signals are crucial in many real-world scenarios:

- **Telemedicine:** Non-contact heart rate monitoring during remote consultations.
- **Structural Vibration Analysis:** Early detection of fatigue in bridges and buildings.
- **Neonatal Monitoring:** Observing subtle breathing patterns in infants without invasive devices.

These examples underscore the practical importance of EVM. Our project extends Wu et al. (2012) by incorporating parameter optimization (via grid search) and face-specific ROI masking to refine signal amplification.

Key Contributions:

- Adaptive frequency and amplification parameter tuning.
- Dynamic ROI masking for targeted amplification.
- Temporal filtering optimization for improved signal fidelity.

2. Problem Statement

Subtle temporal and spatial signals, such as heartbeat-induced color changes, are often hidden in video data. Traditional motion-tracking techniques are computationally expensive and less effective for low-amplitude signals. The goal of this project is to develop an EVM pipeline that:

- Efficiently amplifies imperceptible signals.
- Preserves spatial coherence through multiscale decomposition.
- Minimizes noise and artifacts via adaptive filtering.
- Enables real-time operation without explicit motion estimation.

Formal Definition:

$$E(x,t) = I(x,t) + \alpha\,\delta(t) \cdot \sin\left(\frac{2\pi x}{\lambda}\right)$$

where α is the amplification factor, $\delta(t)$ is local displacement, and λ is the spatial wavelength.

3. Related Work

Wu et al. (2012) laid the foundation for EVM by processing fixed spatial positions, outperforming traditional Lagrangian methods. Poh et al. (2010) provided valuable insights into non-contact cardiac pulse measurements, which motivated our use of dynamic face masking. Liu et al. (2005) further influenced our approach to motion magnification, while additional techniques such as the Laplacian pyramid [?] and optical flow estimation [?] offer complementary methods for capturing motion cues.

3.1. Comparative Analysis

4. Methodology

Study	Approach	Strengths	Weaknesses
Wu et al. (2012)	Eulerian (Laplacian)	Real-time potential, no tracking	Sensitive to noise
Poh et al. (2010)	Cardiac pulse detection	Accurate pulse extraction	Requires stable lighting
Liu et al. (2005)	Lagrangian motion analysis	High motion fidelity	Computationally intensive
This Work	YIQ + ROI + Adaptive EVM	Improved fidelity, reduced noise	1.5× slower runtime

Table 1: Comparative analysis of EVM approaches.

4.1. Basic Implementation

4.1.1 Preprocessing

Video frames are loaded using OpenCV and normalized to [0,1]. All operations occur directly in the RGB space.

4.1.2 Filtering and Reconstruction

Spatial Filtering: Gaussian smoothing.

Temporal Filtering: Basic FFT-based bandpass filtering (0.8–1.0 Hz).

Amplification: Signal amplified by a constant α.

Output: Frames recombined to generate the output video.

Pros: Computationally simple.

Cons: Susceptible to color distortion and background noise.

4.2. Advanced Implementation

4.2.1 Preprocessing & Color Conversion

Color Conversion: Convert from RGB to YIQ to better separate luminance and chrominance, preserving color fidelity.

4.2.2 ROI Masking

Face Detection: Haar Cascade classifiers generate masks for dynamic face tracking.

Mask Resizing: Masks are resized at each pyramid level for consistency.

4.2.3 Multiscale Spatial Decomposition

Laplacian Pyramid: The input frame is decomposed into multiple spatial scales using a Laplacian pyramid.

4.2.4 Temporal Filtering & Adaptive Amplification

Zero-Phase Butterworth Filter: Isolates the 0.8–1.0 Hz frequency bands with minimal phase distortion.

Adaptive Amplification: Amplification is tuned per pyramid level to enhance subtle signals while constraining noise.

Algorithmic Justification: The frequency range of 0.8–1 Hz captures typical heart rate frequencies. The Butterworth filter is chosen for its smooth frequency response and minimal phase distortion.

Amplification Constraint: Ensures amplification does not result in excessive noise or distortion.

4.2.5 Reconstruction

Pyramid Collapse: The decomposed levels are collapsed to restore spatial resolution.

Color Reversion: Convert the processed YIQ frames back to RGB.

Output Generation: Final frames are compiled into the output video.

5. Evaluation

5.1. Dataset

- **Face Videos:** Used for heartbeat detection.
- **Wrist Videos:** Used for pulse detection.
- **Baby Videos:** Used for breathing motion enhancement.
- **Custom Dataset (face3.mp4):** Focuses on subtle facial color changes.

5.2. Experimental Results

Visual Comparisons:

- Clear heartbeat-induced color changes in facial videos.
- Enhanced clarity of wrist pulses and infant breathing patterns.

Quantitative Metrics:

- **SNR Improvement:** 20% increase.
- **SSIM Score:** 0.95.
- **Runtime Efficiency:** The advanced pipeline is approximately 1.5× slower than the basic implementation.

Pipeline	Runtime (s)	Memory (MB)	SNR Improvement (%)
Basic EVM	120	500	10
Advanced EVM	180	650	20

Table 2: Runtime and memory usage comparison between basic and advanced pipelines.

5.3. Enhanced Visual Demonstrations (Video-Based Analysis)

The evaluation relies on comprehensive **video-based demonstrations** rather than static images. All before-and-after videos are consolidated into a single accessible file. These videos clearly show:

- **Before-and-After Comparisons:** Subtle color and motion variations are effectively amplified.
- **ROI Focus:** Dynamic face detection and ROI masking improve amplification clarity while minimizing background noise.

Note: Detailed analyses accompanying these videos are provided in the project notebook.

6. Discussion

The advanced EVM pipeline demonstrates significant improvements over the basic implementation. Minor artifacts in the basic approach—such as slight color distortions and background noise—were mitigated by adaptive amplification and refined filtering. Adjustments in frequency range and amplification parameters were essential for balancing noise reduction with signal enhancement.

Dynamic ROI masking, motivated by Poh et al. (2010), effectively isolated the most relevant regions, enhancing the clarity of amplified signals. Although the advanced pipeline has increased runtime and memory usage (see Table 2), the improvements in accuracy and robustness make it ideal for detailed physiological signal analysis.

7. Conclusion

The advanced Eulerian Video Magnification pipeline effectively amplifies subtle physiological signals and mechanical vibrations, outperforming the basic model in both clarity and robustness. By leveraging adaptive filtering, multiscale decomposition, and targeted ROI amplification, our approach minimizes artifacts while enhancing critical features. These techniques hold significant promise for real-time healthcare applications such as remote monitoring of vital signs, stress analysis, and early detection of cardiovascular conditions. Future work will focus on further runtime optimization and integration with GPU-based processing to facilitate real-time implementations.

8. 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] Liu, C., Torralba, A., Freeman, W. T., Durand, F., & Adelson, E. H. (2005). *Motion Magnification.* ACM Transactions on Graphics.

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

[4] Burt, P., & Adelson, E. (1983). *The Laplacian Pyramid as a Compact Image Code.* IEEE Transactions on Communications.

[5] Horn, B., & Schunck, B. (1981). *Determining Optical Flow.* Artificial Intelligence.