Chapter 14
Non-contact Video-based Heart Rate Monitoring using Hilbert Motion Magnification

Camilly S. Costa, Gustavo N. Castro, Yuri S. A. Alencar, Victor H. R. Cardoso, João C. W. A. Costa and Moisés F. Silva

Abstract Heart rate measurement is vital for health monitoring and medical diagnosis. Although traditional methods like oximeters and electrocardiograms are precise, they have limitations such as the need for physical contact, high costs, and operational complexity. Therefore, this study explores a non-invasive, cost-effective alternative: measuring subtle head movements caused by the Newtonian blood flow reaction from heartbeats captured via video measurements. The process involves applying the Hilbert transform to obtain the analytical signal, from which we extract its real and imaginary components. Next, Principal Component Analysis is employed to compress the data by identifying the principal components from each set of real and imaginary data. Subsequently, the identified principal components are combined into a single dataset for further analysis. At this point, we apply blind source separation using complexity pursuit to isolate individual motion patterns that closely match the heartbeat responses measured by the oximeter. Tests with individuals in various situations have shown the achieved results were nearly identical to those obtained by a reference oximeter measuring real-time heart rate. We also establish comparisons to a widely known phase-based optical flow algorithm on benchmark videos. It is demonstrated that the proposed approach can derive comparable results in a simple fashion, reducing the number of steps for data processing and the efforts in terms of setup preparation. The study indicates the proposed video-based heart rate estimation approach is a viable and effective alternative, offering comparable results to conventional and other video-based alternatives. This non-invasive approach can be widely applied in medical fields and health monitoring contexts, significantly advancing the democratization of diagnostic and monitoring tools. It has the potential to be optimized for system embedding, such as into smartphones and security cameras, to ease its integration in practical applications.

Keywords Hilbert transform · Non-contact monitoring · Heart rate · Blind source separation · Principal components analysis

Introduction

Heart rate (HR) measurement is an essential practice for assessing cardiovascular health and monitoring in various clinical situations. Its importance extends to areas such as sports medicine, rehabilitation, and stress assessment. HR is a sensitive indicator of the body's physiological responses to different stimuli, such as physical exercise, mental stress, and pathological conditions [1]. Heart rate variability is a relevant indicator, as it reflects the ability of the autonomic nervous system to regulate HR in response to various conditions. High variability is generally associated with good cardiovascular health and an efficient adaptive response to stress, while low variability may indicate autonomic dysfunction and is linked to pathological states, such as specific stress and cardiovascular diseases. To further elaborate, studies show that regular monitoring of heart rate can help in the early detection of abnormalities in cardiovascular function, thereby preventing severe conditions such as myocardial infarction and stroke. This underscores the importance of accessible and accurate heart rate measurement tools in both clinical and remote health monitoring contexts. [2]

Camilly S. Costa · Gustavo N. Castro · Yuri S. A. Alencar · Victor H. R. Cardoso · João C. W. A. Costa
Laboratório de Eletromagnetismo Aplicado, Universidade Federal do Pará, Belém, PA 66075, Brazil
e-mail: camilly.costa@itec.ufpa.br; gustavonascimentocastro1@gmail.com; yuri.alencar@itec.ufpa.br; victorcard@ufpa.br; jweyl@ufpa.br

Moisés F. Silva
Los Alamos National Laboratory, Los Alamos, NM 87545
e-mail: mfelipe@lanl.gov

Matthew Allen et al. (eds.), *Special Topics in Structural Dynamics & Experimental Techniques, Vol. 5*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/97887-438-0150-4_14

The demand for telemedicine services and contactless medical monitoring methods has increased significantly, especially in response to the COVID-19 pandemic. Telemedicine, which includes virtual consultations and remote monitoring, has emerged as a viable solution to ensure continuity of medical care during periods of social restrictions and increased demand for healthcare. The use of information and communication technologies has been crucial to the expansion of these services, enabling safe and efficient interactions between doctors and patients [3]. Studies show that telemedicine not only facilitates access to care but also improves patient satisfaction. The implementation of telehealth services, such as teleconsultations and telemonitoring, has shown positive results in terms of clinical control and treatment adherence, especially in vulnerable populations, such as the elderly and those with chronic diseases [4].

Video-based heart rate (HR) monitoring has gained attention for its non-invasive nature and convenience, with optical flow-based algorithms standing out due to their ability to detect blood pulses by analyzing variations in light reflected by the skin. These algorithms, alongside techniques like Eulerian motion magnification, allow video cameras to capture subtle changes caused by blood flow, enabling remote HR estimation without the discomfort of electrodes [5, 6]. Despite their benefits, such methods face limitations, particularly when influenced by environmental factors like lighting variations and patient movement. The accuracy of these techniques heavily depends on image quality and the algorithm's ability to process data in real-time, with ambient light fluctuations sometimes leading to inaccurate measurements.

Proposed [7], the Hilbert transform-based video motion magnification method offers significant advantages over other techniques, such as Eulerian motion magnification. While Eulerian motion magnification is effective in revealing subtle temporal variations in videos, it faces challenges related to noise sensitivity and the need for a controlled environment to obtain reliable results. In contrast, the Hilbert transform-based approach allows for more robust handling of small motions by utilizing the imaginary component of the transform to access the amplitude of phase variations in the video over time. This technique is less susceptible to external interference and can withstand higher amplification factors, making it more suitable for applications in uncontrolled environments.

Given this, the aim of this paper is to explore a novel, non-invasive and cost-effective approach for HR monitoring using Hilbert-based motion magnification. The proposed methodology involves capturing subtle head movements resulting from the reaction of blood flow to heartbeats through video measurements. Initially, the Hilbert transform is applied to obtain the analytical signal, from which its real and imaginary components are extracted. Tests conducted with individuals in various situations demonstrated that the results obtained with the proposed approach were almost identical to those measured by a real-time reference oximeter.

The main contributions of this work include simplifying the process of non-contact HR monitoring and proving its effectiveness in practical situations. This non-invasive approach can be widely applied in medical and health monitoring contexts, significantly advancing the democratization of diagnostic and monitoring tools. Furthermore, the method has the potential to be optimized for incorporation into embedded systems, such as smartphones and security cameras, facilitating its integration into practical and everyday applications.

Theory

To measure heart rate (HR) by extracting motion patterns, we start from the premise that the algorithm detects a person's pulse by analyzing the periodic movement of the head caused by the pulsation of blood in the carotid arteries. This phenomenon, often referred to as Newtonian motion, results from the mechanical forces of blood flow during each heartbeat, based on Newton's third law of motion. As blood is pumped from the heart to the head, primarily through the carotid arteries, it generates subtle but measurable movements of the head due to the forces involved in this pulsate flow.[8, 9]

This displacements can be captured through video recordings that show frames with displaced or time-shifted image intensity. As described in [10], the algorithm is based on a combination of spatial and temporal processing to emphasize subtle changes in a video. After processing, the video goes through three stages to obtain the necessary parameters and calculate the heart rate: Hilbert Transform(HT), Principal Component Analysis (PCA), and Blind Source Separation (BSS). Once processed, the video is reconstructed with the appropriate amplification factor.

The algorithm combines the Hilbert Transform, PCA, and BSS to analyze the time series of pixels obtained from video measurements. The process begins by transforming the pixel intensity measurements into a displacement representation over time using the Hilbert Transform. This approach extracts the real and imaginary parts of the analytic signal estimated from the video, providing motion information of the scene.

The video is then reshaped into a matrix or time series, allowing for further processing. The first step is applying the Hilbert Transform. The HT is defined as the convolution of a function in the frequency domain. This phase shift implies that the imaginary part of the transformed signal is related to the real part through a specific phase relationship [11]. By calculating the instantaneous frequency and amplitude of signal components, it is possible to visualize the movement of objects in the scene. The resulting spectral analysis provides detailed information on how objects move over time.

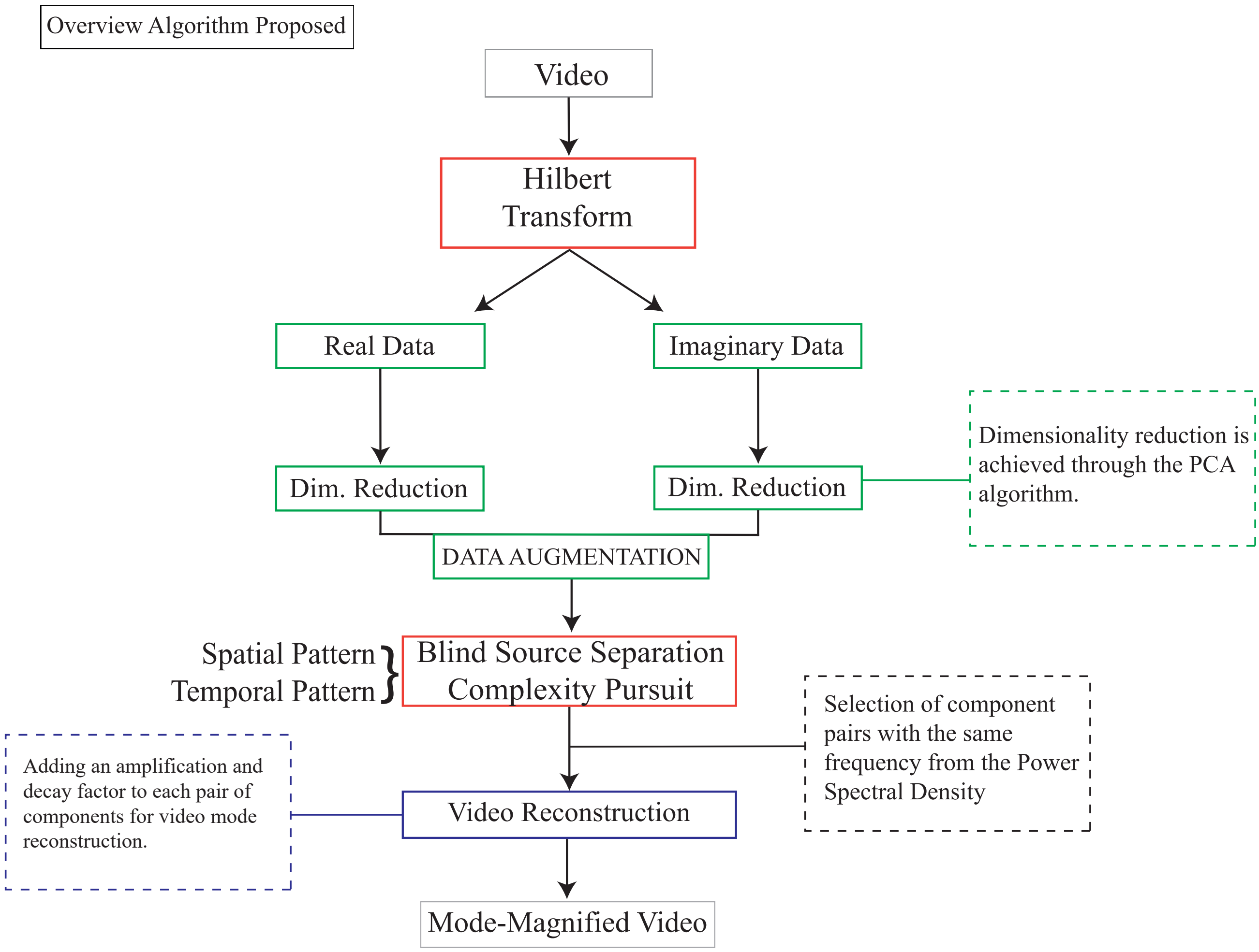

Fig. 1 Overview of the video Hilbert magnification algorithm.

Principal components analysis

After extracting the real and imaginary parts of the pixel time series, Principal Component Analysis is applied to reduce the dimensionality of these matrices. PCA is a widely recognized and used statistical technique for dimensionality reduction, especially in high-dimensional datasets like videos. Its main goal is to transform the original dataset into a new set of orthogonal components that preserve most of the variance present in the original data, but with fewer dimensions. Mathematically, PCA starts by calculating the covariance matrix C of the data. Where X is the centered data matrix, and n is the number of observations. Once the covariance matrix is obtained, the eigenvalues λ and eigenvectors v are calculated as follows:

$$Cv = \lambda v \tag{1}$$

The principal components are then determined by ordering the eigenvalues in descending order and selecting the first k eigenvectors, where k is the desired number of dimensions. This process reduces the dimensionality of the dataset, preserving the most significant characteristics, which is crucial for subsequent analysis [12]. This allows for the elimination of redundancies and noise, improving the accuracy of subsequent steps, such as BSS.

Blind source separation

After applying PCA, the reduced data is processed using Blind Source Separation (BSS). BSS is widely used for the modal identification of individual motion patterns. Here, the Complexity Pursuit (CP) is used to identify motion parameters of structures subjected to ambient excitation. These algorithms belong to the BSS class, which aims to recover mixed sources from their observed outputs without prior knowledge of the mixing process or original sources. The relationship between the mixed signals *x(t)* and the sources *s(t)* can be mathematically represented as:

$$x(t) = As(t) \tag{2}$$

Where A is the unknown mixing matrix, and $s(t)$ are the independent source signals. The goal of Blind Source Separation is to find the matrix A and the sources $s(t)$, using only the observations $x(t)$. Once the mixing matrix A is estimated, the sources can be recovered by inverting the matrix

$$s(t) = A^{-1}x(t) \tag{3}$$

Complexity pursuit

In the Complexity Pursuit algorithm [13], the temporal complexity of the extracted signals is minimized. Complexity is defined as the unpredictability of a signal over time, and CP seeks to find a linear combination of the mixed signals that is the least complex, or most predictable. The complexity function $F(w_i, x)$ to be minimized is:

$$F(w_i, x) = \frac{V(w_i, x)}{U(w_i, x)}$$

Where $V(w_i, x)$ represents the long-term variance of the extracted signal, and $U(w_i, x)$ represents the short-term variance. The separation matrix W is found by maximizing the temporal predictability of the separated signals. The resulting signal sources correspond to the temporal patterns of each mode of movement in the scene. Based on this, the columns of the mixing matrix provided by BSS are used to expand the identified motion patterns to all frames of the video at high spatial resolution, using the projection matrix obtained during the PCA stage. This process allows the detected movement modes to be applied to the entire visual content, resulting in a detailed and accurate representation of the subtle movements present in the scene. After BSS, temporal and spatial patterns can be extracted, from which we can identify the temporal frequency of the subtle displacements caused by the heart rate pattern.

One of the main advantages of the proposed algorithm over Eulerian Motion Magnification is the absence of the need for manual selection of a region of interest (ROI) in the video. While the Eulerian method requires the selection of a specific area to amplify movements, the algorithm automatically amplifies movements across the entire scene. This eliminates selection bias, enabling a more complete and automated analysis of motion.

Moreover, the use of BSS in the algorithm allows for the isolation of relevant movements without amplifying noise or other unwanted movements. In the Eulerian method, all movements are amplified globally, which can lead to noise amplification and distortions, especially in noisy videos. Additionally, unlike phase-based optical flow, this method does not require the manual creation of spatial filters tailored to the movement direction. For example, if the displacements are horizontal, a horizontal filter is needed for optical flow, which requires prior knowledge of the motion direction in the video. This necessitates direct and critical human intervention. By separating and analyzing motion sources individually, the method results in more precise and focused amplification without the need for such manual adjustments.

Experimental results

The approach was implemented and executed in Python(3.12.16 version). Was performed on a device featuring an 11th Gen Intel(R) Core(TM) i5-1135G7 processor running at 2.40 GHz, with 24 GB of installed RAM. The system is a 64-bit operating system with an x64-based processor. The videos were recorded using mobile phone cameras in non-isolated environments with varying lighting conditions. All videos have a frame rate of 60 frames per second in HD resolution. Oximeters were attached to the subjects for comparison of the results.The results are presented in pairs because, at the beginning of the process, the Hilbert transform outputs the data in two components: one representing the real part and the other the imaginary part.

After identifying the principal frequency, we apply a simple formula to convert the frequency from Hertz (Hz) to beats per minute (BPM), where f is the frequency in Hertz:

$$\text{BPM} = 60 \times f \tag{4}$$

In Figures 3, 4, 5, and 6, the tests were conducted with individuals facing the camera. As shown in Table 1, the error between the results obtained from the oximeter and the frequency extracted by the algorithm is minimal, further demonstrating the effectiveness of the technique.

Visual representation of data from each test scenario

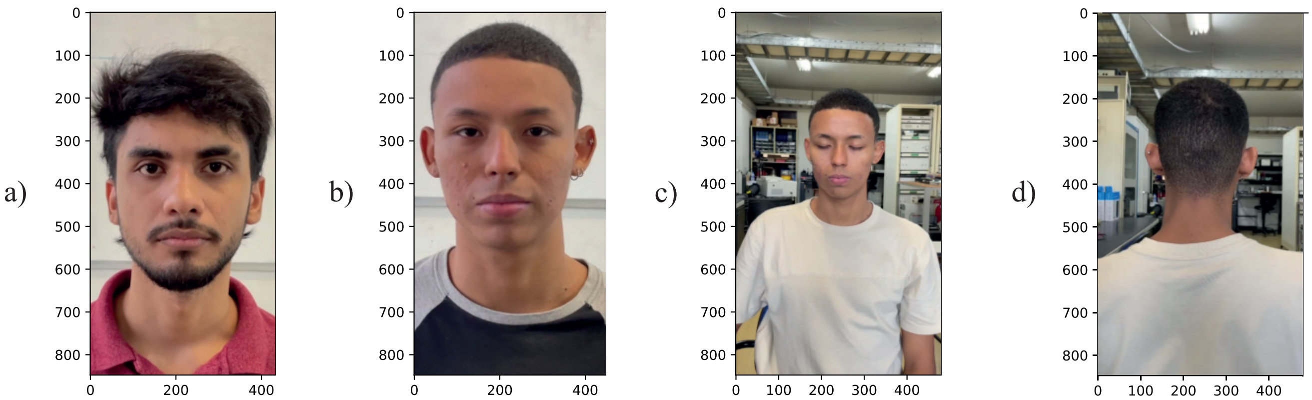

Fig. 2 (a) Subject 1, (b) subject 2, (c) subject's eyes closed, and (d) subject facing away.

Up to this point, promising results have been obtained. However, to push the analysis further, additional tests were conducted from different perspectives. A test with closed eyes (Figures 5 and 6) was performed to determine whether blinking could influence the outcome. Additionally, another test was carried out with the head positioned at a different angle, facing away from the camera (Figures 7 and 8), to assess where head movements related to arterial pulsation could be more effectively captured by the algorithm.

In all experiments, the error was minimal, and variations in head angle did not significantly impact the results. Each setup successfully captured the subtle oscillations of the head caused by the pulsation of the aorta artery, and the heart rate was consistently detected with accuracy.

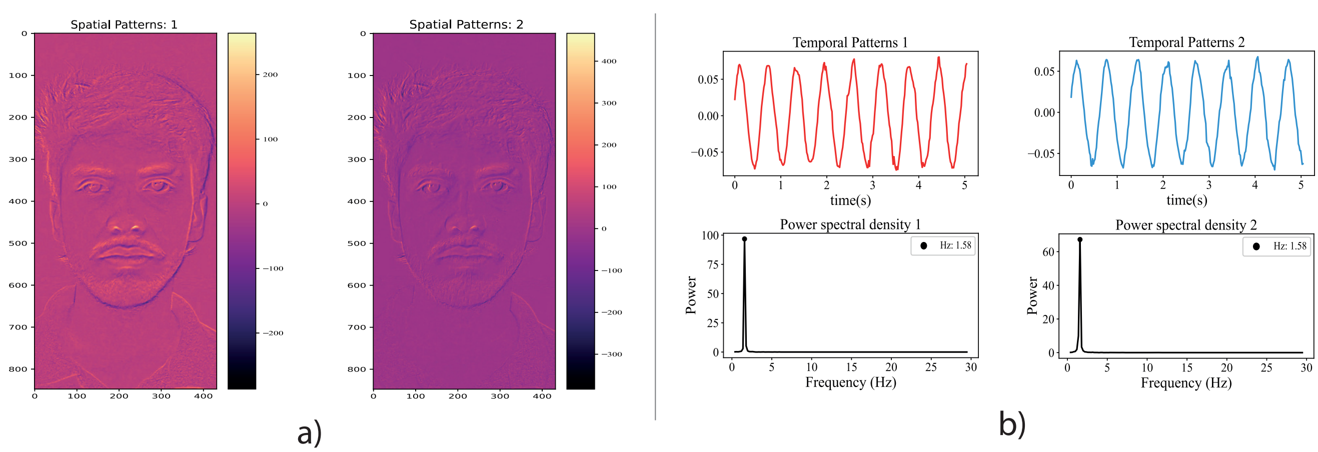

Fig. 3 (a) Spatial patterns, (b) temporal patterns, and power spectral density from the test with subject 1.

Table 1 Comparison of Test Results with Oximeter Readings in BPM

	Subject 1	Subject 2	Eyes closed	Facing away
Oximeter	100	73	75	75
Hilbert Magnification	94.8	72	72.6	74.4
Error(%)	5.2	1.37	3.2	0.8

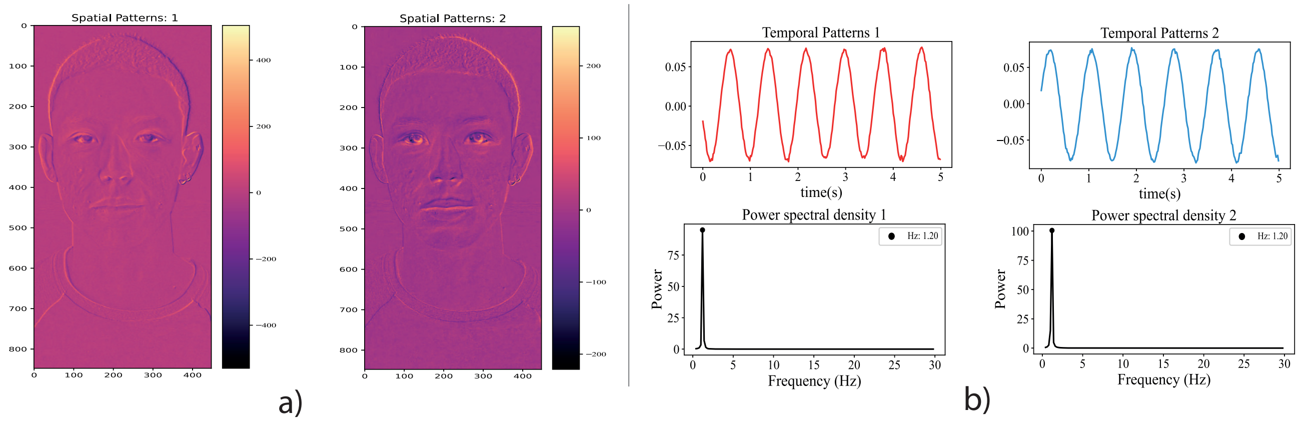

Fig. 4 (a) Spatial patterns, (b) temporal patterns, and power spectral density from the test with subject 2.

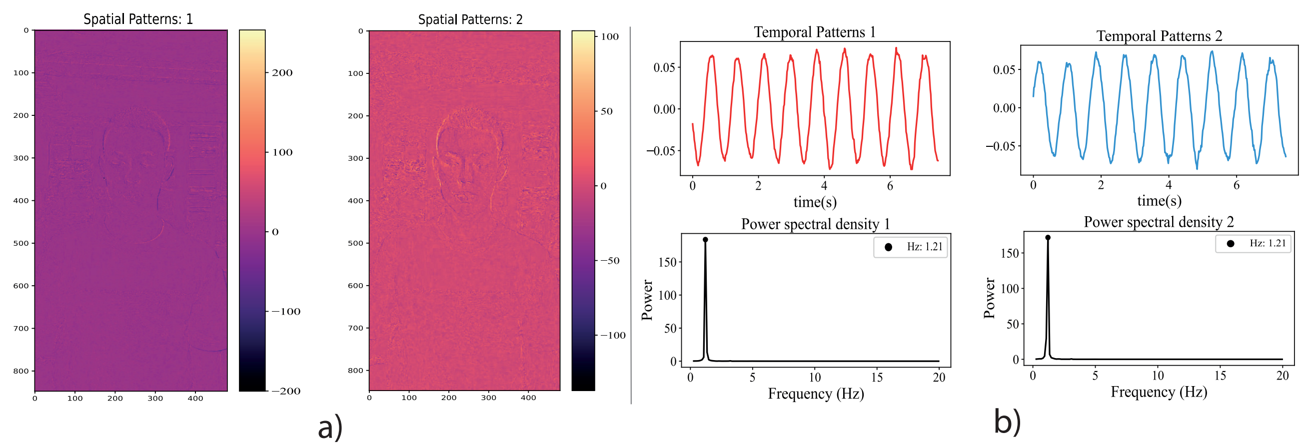

Fig. 5 (a) Spatial patterns, (b) temporal patterns, and power spectral density from the test with subject with the subject's eyes closed.

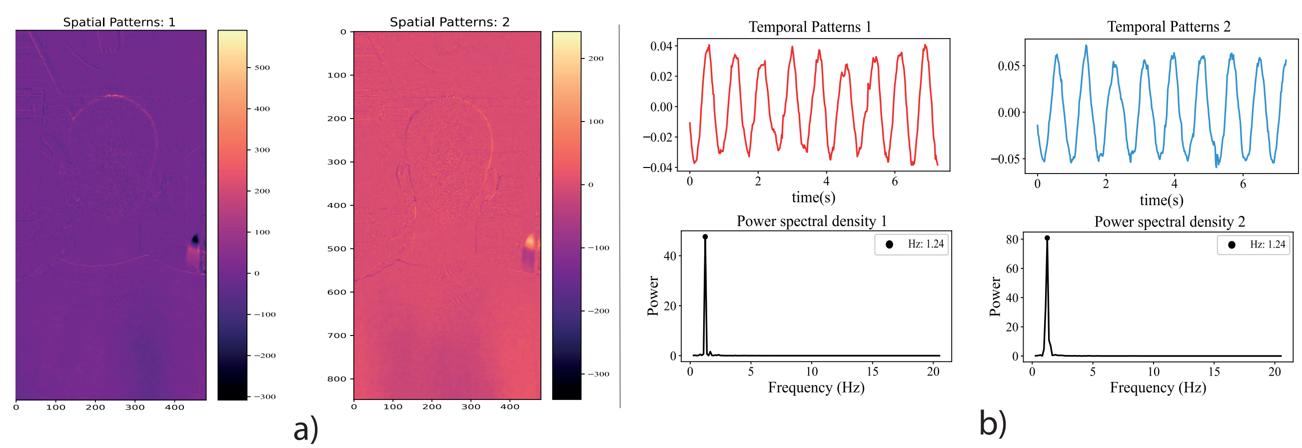

Fig. 6 (a) Spatial patterns, (b) temporal patterns, and power spectral density from the test with subject with the subject facing away.

Conclusion

This study explored an innovative, non-invasive video-based heart rate monitoring approach using Hilbert motion magnification. The proposed methodology proved to be effective in capturing subtle head movements caused by blood pulsation, providing results comparable to traditional methods such as oximeters. Furthermore, comparisons with other algorithms demonstrated that the developed technique can achieve similar accuracy while simplifying the data processing steps and configuration preparation.

The results of the tests, carried out under various lighting conditions and head angles, indicate that the algorithm can be widely applied in health monitoring contexts without the need for physical contact, which is advantageous for applications such as telemedicine and remote monitoring. With potential for optimization and integration into embedded systems, such as smartphones and security cameras, the presented technique offers an accessible and low-cost solution for monitoring cardiovascular health. This approach can significantly contribute to the democratization of diagnostic and monitoring tools, paving the way for the development of technologies that promote more inclusive and efficient healthcare.

Acknowledgments It was supported by the Laboratory of Applied Electromagnetism (LEA) at the Federal University of Pará. Furthermore, the authors acknowledge that this study was financed in part by the Amazon Foundation to Support Studies and Research (FAPESPA), Grant agreement 050/2023. It was also financed by the São Paulo Research Foundation (FAPESP) with the support provided by grant 2022/10105-3, as well as the National Council for Scientific and Technological Development (CNPq).

References

1. Kim, H.G., Cheon, E.J., Bai, D.S., Lee, Y.H., and Koo, B.H. "Stress and heart rate variability: a meta-analysis and review of the literature". *Psychiatry investigation*, 15(3):235 (2018)
2. Tiwari, R., Kumar, R., Malik, S., Raj, T., and Kumar, P. "Analysis of heart rate variability and implication of different factors on heart rate variability". *Current cardiology reviews*, 17(5) (2021)
3. Lee, S.M. and Lee, D. "Opportunities and challenges for contactless healthcare services in the post-covid-19 era". *Technological Forecasting and Social Change*, 167:120712 (2021)
4. Silva, M. et al. "Telemedicine during covid-19 pandemic". *Journal of Health Informatics*, 15:33–40 (2021)
5. Zhu, Q., Wong, C.W., Fu, C.H., and Wu, M. "Fitness heart rate measurement using face videos". In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 2000–2004. IEEE (2017)
6. Wu, H.Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., and Freeman, W. "Eulerian video magnification for revealing subtle changes in the world". *ACM transactions on graphics (TOG)*, 31(4):1–8 (2012)
7. Silva, M., Green, A., and Mascareñas, D. "A generalized technique for full-field blind identification of travelling waves and complex modes from video measurements with hilbert transform". In *Data Science in Engineering, Volume 9: Proceedings of the 39th IMAC, A Conference and Exposition on Structural Dynamics 2021*, pages 233–236. Springer (2022)
8. Kaid, N., Benyamina, L., Menni, Y., Alkhafaji, M.A., Bayram, M., Alshammari, B.M., and Kolsi, L. "Unveiling hemodynamic pulsatile flow dynamics in carotid artery stenosis: Insights from computational fluid dynamics". *AIP Advances*, 14(6):065128 (2024)
9. Hammoud, A., Sharay, E.Y., and Tikhomirov, A.N. "Newtonian and non-newtonian pulsatile flows through carotid artery bifurcation based on ct image geometry". In *AIP Conference Proceedings*, volume 2171, page 110022 (2019)
10. Mayorga, P., Rivoldini, A., and Rixen, D.J. "Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification". *Mechanical Systems and Signal Processing*, 85:567–590 (2017)
11. Liu, G., Li, M., Mao, Z., and Yang, Q. "Structural motion estimation via hilbert transform enhanced phase-based video processing". *Mechanical Systems and Signal Processing*, 166:108418 (2022)
12. Jolliffe, I.T. and Cadima, J. "Principal component analysis: a review and recent developments". *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202 (2016)
13. Hyvärinen, A. "Complexity pursuit: separating interesting components from time series". *Neural computation*, 13(4):883–898 (2001)