

Heart Rate Detection using Video Magnification: Impact of Algorithmic Parameters and Noise

by

Leen Yassin Kassab

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in
partial fulfillment of the requirements for the degree of

Master of Applied Science

in

Electrical and Computer Engineering

Carleton University
Ottawa, Ontario

© 2022

Leen Yassin Kassab

Abstract

Screening people for illnesses in public spaces using contactless techniques such as Video Magnification (VM) can be beneficial for both the crowd and the operators. One important vital sign to screen for is heart rate because it is indicative of the overall health of humans. In this thesis, first the impact of varying two input parameters (region of interest (ROI) size and window segment length) on VM is explored. Second, the impact of noise from three different sources (quantization, modern camera systems coupled with software enhancements, and light illumination) is explored. Results show that heart rate can be detected from very small ROIs, but that larger ROIs have higher accuracy. Larger window segment lengths are useful in the presence of motion. VM detects signals with amplitudes equal to one quantization level in 8-bit videos. VM is affected by modern camera systems, and it is more accurate with higher light illuminations.

Acknowledgements

The work in this thesis was supported by the National Research Council of Canada's Pandemic Response and Aging in Place Challenge Programs and NSERC. This work was also supported in part by the AGE-WELL NCE Inc.

I would like to thank my supervisors: Dr. Rafik Goubran and Dr. Bruce Wallace for their great support and guidance. I would also like to thank Dr. Andrew Law, the primary researcher at the National Research Council, Dr. Frank Knoefel, MD at Bruyere hospital, and Julien Lariviere-Chartier, the lab technician for their valuable contributions to my research.

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	vii
List of Figures	viii
List of abbreviations:	xii
Chapter 1: Introduction	1
1.1 Problem Statement	1
1.2 Objectives	2
1.3 Summary of Contributions.....	2
1.4 Thesis Organization	4
Chapter 2: Background Review	5
2.1 Heart Rate Signal from a Clinical Perspective.....	5
2.2 State of the Literature on the Video Magnification Technique and its Limitations.....	6
2.2.1 Summary of Review Papers on Remote Heart Rate Assessment	6
2.2.2 Limitations of Contactless Heart Rate Measurement through Video	8
2.3 Work to Address the Limitations of Video Magnification	10
Chapter 3: Method and Experimental Setup	12
3.1 Video Capture:	12
3.1.1 Video Capture Method from the National Research Council Collaboration.....	12
3.1.2 Self-Captured Videos Method	14
3.1.3 Artificially Generated Videos in MATLAB.....	16
3.1.3.1 Video Generation without Noise	16
3.1.3.2 Video Generation with Noise.....	18

3.2	Ground-truth Heart Rate Capture.....	19
3.2.1	Ground-truth for the NRC Video Capture.....	19
3.2.2	Ground-truth for the Self-Captured Videos.....	20
3.3	Video Analysis Method	20
Chapter 4: Effects of Algorithm Parameters.....		23
4.1	Effect of Region of Interest Size on the Video Magnification Algorithm.....	23
4.1.1	Analysis Details.....	23
4.1.2	Effect of Small Region of Interest Size:.....	26
4.1.3	Effect of Large Region of Interest Size	28
4.1.4	Comparing Small and Large ROI Sizes.....	31
4.1.5	Continuous Absolute Error for the Four Conditions	31
4.2	Effect of Window Segment Length on the Video Magnification Algorithm.....	35
4.2.1	Analysis Details.....	36
4.2.2	Effect of Three Window Segment Length: 10s, 20s, 30s.....	37
4.3	Conclusion	47
Chapter 5: Effect of Noise		48
5.1	Effect of Quantization Noise on the Video Magnification Algorithm.....	49
5.1.1	Artificial Video Processing	50
5.1.2	Effect of Large and Small Signal Amplitude in the Absence of Noise	50
5.1.3	Effect of Small Signal Amplitude in the Presence of Noise.....	58
5.2	Effect of Camera on the Video Magnification Algorithm.....	60
5.2.1	Video Magnification Performance with VGA Web Camera.....	61
5.2.2	Video Magnification Performance with iPhone 12 Pro Max Camera	62
5.3	Effect of Illumination Level on the Video Magnification Algorithm.....	66
5.3.1	Analysis Details.....	66
5.3.2	Effect of Two Illumination Levels from the Combination of Two Light Sources	68

Chapter 6: Conclusions.....	74
6.1 Conclusion of Thesis Contributions.....	74
6.2 Guidelines for Video Capture and Optimal Video Magnification Results	76
6.3 Suggestions for Future work.....	76
References	78

List of Tables

Table 1. Summary of the experimental conditions for the tests done by NRC.....	14
Table 2. Smart camera system iPhone 12 Pro Max.....	15
Table 3. Summary of the smart video features in iPhone 12 Pro Max	16
Table 4. A summary of the four testing conditions used in the analysis [10].....	24
Table 5. Summary of the percent mean absolute error of the 19 subjects for three window segment lengths: 10s, 20s and 30s.	38
Table 6. Average, standard deviation and mean of the percent mean absolute error of the 19 subjects for the three window segment lengths: 10s, 20s and 30s.	39
Table 7. Summary of the two conditions used in the analysis	67
Table 8. % mean absolute error results for the two lighting configurations from [11].....	69
Table 9. Average, standard deviation and maximum % mean absolute error for the 19 subjects from [11]	69
Table 10. Comparison of the percent mean absolute error for the four conditions (C1-C4) using small and large ROI from [10]	74

List of Figures

Figure 1. Illustration of the video generation method.....	17
Figure 2. Illustration of the video generation method with noise obtained from [12]	19
Figure 3. Inputs and outputs of the Video Magnification algorithm using a video frame of the author.....	21
Figure 4. Flowchart containing the detailed steps of the Video Magnification algorithm.....	22
Figure 5. Areas covered by the different large ROI sizes: top left: 16x16, bottom left: 32x32, top right: 64x64, bottom right: 256x256, middle center: 512x512 from [10].....	25
Figure 6. Relationship between the small ROI and the large ROI analysis (small 8x8, large 32x32).....	25
Figure 7. % mean absolute error for average of the 16 contiguous small ROIs- green from [10]	27
Figure 8. % mean absolute error for average of the 16 contiguous small ROIs - red from [10] ..	27
Figure 9. % mean absolute error for average of the 16 contiguous small ROIs - blue from [10].	28
Figure 10. % mean absolute error for large ROIs - green channel from [10]	29
Figure 11. % mean absolute error for large ROIs - red channel from [10].....	29
Figure 12. % mean absolute error for large ROIs - blue channel from [10]	30
Figure 13. Continuous % mean absolute error of C1,C4,C5,C8, green channel, ROI 256x256 from [10]	33
Figure 14. VM algorithm results for condition C1 (sitting, no mask), green, 256x256 ROI, from [10]	33
Figure 15. VM algorithm results for condition C4 (sitting, with mask), green, 256x256 ROI from [10]	34

Figure 16. VM algorithm results for condition C5 (standing, no mask), green, 256x256 ROI from [10]	34
Figure 17. VM algorithm results for condition C8 (standing, with mask), green, 256x256 ROI from [10]	35
Figure 18. The VM algorithm results for subject S4 using a 10s window segment length	39
Figure 19. The VM algorithm results for subject S18 using a 10s window segment length	40
Figure 20. The VM algorithm results for subject S11 using a 10s window segment length	40
Figure 21. The VM algorithm results for subject S11 using a 20s window segment length	41
Figure 22. The VM algorithm results for subject S11 using a 30s window segment length	41
Figure 23. The VM algorithm results for subject S19 using a 10s window segment length	43
Figure 24. The VM algorithm results for subject S19 using a 20s window segment length	43
Figure 25. The VM algorithm results for subject S19 using a 30s window segment length	44
Figure 26. True heart rate from PPG for subject S14.....	45
Figure 27. The VM algorithm results for subject S14 using a 10s window segment length	45
Figure 28. The VM algorithm results for subject S14 using a 20s window segment length	46
Figure 29. The VM algorithm results for subject S14 using a 30s window segment length	46
Figure 30. Sources of noise in remote heart rate measurement from video	48
Figure 31. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.05 .	51
Figure 32. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.04 .	52
Figure 33. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.03 .	52
Figure 34. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.02 .	53
Figure 35. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.01 .	53

Figure 36. Average of the peak FFT magnitude at the true heart rate for $k = 0.01-0.05$ from [12] 54

Figure 37. VM algorithm results for the 53BPM artificial video at small signal amplitude 0.005 55

Figure 38. VM algorithm results for the 53BPM artificial video at small signal amplitude 0.004 56

Figure 39. VM algorithm results for the 53BPM artificial video at small signal amplitude 0.003 from [12] 56

Figure 40. VM algorithm results for the 53BPM artificial video at small signal amplitude 0.002 from [12] 57

Figure 41. VM algorithm results for the 53BPM artificial video at signal amplitude 0.00196 57

Figure 42. Average of the peak FFT magnitude at the true heart rate for $k = 0.002-0.005$ from [38] 58

Figure 43. Average and standard deviation of the FFT Ratio for the different noise levels m from [12] 59

Figure 44. VM algorithm results for the 53BPM artificial video at $k= 0.002$ and $m=15$ from [38] 60

Figure 45. The VM algorithm results for a VGA camera video 61

Figure 46. True heart rate results from PPG for VGA video 62

Figure 47. The VM algorithm results for an iPhone 12 Pro Max video with all smart features turned off..... 63

Figure 48. The VM algorithm results for an iPhone 12 Pro Max video with smart HDR feature turned on..... 63

Figure 49. The VM algorithm results for an iPhone 12 Pro Max video with smart Auto 30FPS turned on..... 64

Figure 50. The VM algorithm results for an iPhone 12 Pro Max video with smart Auto 30&60 FPS turned on 64

Figure 51. The VM algorithm results for an iPhone 12 Pro Max video with smart Auto 30&60FPS and HDR turned on..... 65

Figure 52. True heart rate results from PPG for iPhone 12 Pro Max videos 65

Figure 53. Illumination levels of the two lighting conditions C1 and C2 obtained from [11]..... 67

Figure 54. The VM algorithm results for condition C1 subject S4 (yawning) 71

Figure 55. The VM algorithm results for condition C2 subject S4..... 71

Figure 56. True Heart Rate Results from PPG Subject S4..... 72

Figure 57. VM algorithm results for subject S11 condition C2 obtained from [11]..... 73

Figure 58. True heart rate from PPG for subject S11 condition C2 obtained from [11]..... 73

List of abbreviations:

EVM: Eulerian Video Magnification

VM: Video Magnification

HR: Heart Rate

ROI: Region of Interest

NRC: National Research Council

r-PPG: Remote Photoplethysmography

PPG: Photoplethysmography

RGB: Red/Green/Blue

IR: Infrared

FFT: Fast Fourier Transform

FPS: Frames Per Second

BPM: Beats Per Minute

Chapter 1: Introduction

This chapter first covers the problem statement of this thesis followed by the objectives, the summary of contributions and finally the thesis structure.

1.1 Problem Statement

Vital signs that are comprised of heart rate (HR), body temperature, blood pressure and respiration rate give an indication of the overall health status of a human being. These vitals are frequently the first things that get checked during a visit to the family doctor in clinic or the emergency department at the hospital. Heart rate is an important indicator of the health status of the heart organ responsible for circulating blood within the body. Tracking heart rate changes over time can give an insight into the human's mental state as well as physical wellbeing [1]. Additionally, with age, monitoring heart rate becomes more important for early detection of illnesses and diseases including hypertension and diabetes [1].

Methods to assess heart rate include invasive (placement of sensors inside the body) and non-invasive techniques. The non-invasive techniques are further divided into contact methods (placing sensors on the skin) [2], [3] and contactless methods (video camera, microphone) [4]. Over the last decade, advances had been made in the area of video based contactless remote assessment of heart rate. These advances began with the publication of two important works: Eulerian Video Magnification (EVM) [5] and Remote Photoplethysmography (rPPG) [6] in 2012 and 2013 respectively.

Contactless techniques for measuring heart rate such as Video Magnification are beneficial for screening people for early signs of illnesses in public spaces without obstructing their activities [7]. They are also useful for screening frailer people in places such as long-term care facilities [8]. However, the Video Magnification technique has limitations in terms of the correct choices of

algorithmic parameters as well as its sensitivity to noise.

1.2 Objectives

The objective of this thesis is to improve the accuracy of the Video Magnification algorithm at detecting heart rate. This is achieved through first investigating two of the algorithm's input parameters and improving the choices for these parameters and then by investigating three sources of noise and quantifying their effects on the algorithm. The two input parameters are the region of interest (ROI) and the window segment length that are used for the measurement of heart rate from facial videos. The three sources of noise are quantization noise, noise from modern camera systems equipped with software enhancement features, and noise from illumination levels from different light sources.

1.3 Summary of Contributions

This thesis is a summary of the research performed over the last year and a half in collaboration with the National Research Council's (NRC) Aerospace team as part of the Pandemic Response Challenge and Aging in Place programs. In addition to myself, the research group was comprised of Dr. Andrew Law, the lead researcher from NRC, Julien Larivière-Chartier, a research technologist, Dr. Frank Knoefel, a medical doctor, and my two supervisors: Dr. Bruce Wallace and Dr. Rafik Goubran. The research resulted in the following contributions:

1. As lead author, a report on the current literature for contactless heart rate and respiration rate techniques delivered to NRC.

[9] L. Yassin Kassab, B. Wallace, R. Goubran, and F. Knoefel, "Video Magnification Remote Sensing of Heart Rate and Respiration Rate," Carleton University Report for the National Research Council of Canada, 2021.

The research team then took the opportunity to rework the NRC report into an

Instrumentation and Measurement Magazine article and Dr. Wallace took the lead on it.

[4] B. Wallace, L. Y. Kassab, A. Law, R. Goubran, and F. Knoefel, "Contactless Remote Assessment of Heart Rate and Respiration Rate Using Video Magnification," *IEEE Instrum. Meas. Mag.*, vol. 25, no. 1, pp. 20–27, Feb. 2022

A key contribution of the two mentioned works was the knowledge that the two different threads of work termed remote-PPG and Eulerian Video Magnification although presented independently are now overlapping and should be considered as one combined body of work.

2. Better knowledge and methods around the region of interest (ROI) selection presented in an IEEE conference paper using analysis results of four videos of the NRC researcher taken as part of a larger multi-modal contactless vital sign assessment solution developed at NRC.

[10] L. Yassin Kassab, A. Law, B. Wallace, J. Larivière-Chartier, R. Goubran, and F. Knoefel, "Effects of Region of Interest Size on Heart Rate Assessment through Video Magnification," in *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2021, p. 6.

3. Better understanding of the effect of the segment duration length on the Video Magnification algorithm presented in an IEEE conference paper.

[11] L. Yassin Kassab, A. Law, B. Wallace, J. Larivière-Chartier, R. Goubran, and F. Knoefel, "Effects of Lighting and Window Length on Heart Rate Assessment through Video Magnification," 2022, p. 6. In review for *2022 IEEE Sensors Applications Symposium Conference (SAS)*, 2022.

4. Better understanding of the Video Magnification algorithm's sensitivity to noise and the effects/benefits of small and large White Gaussian noise in the video presented in an IEEE

conference paper.

[12] L. Yassin Kassab, A. Law, B. Wallace, J. Larivière-Chartier, R. Goubran, and F. Knoefel, “The Effect of Noise on Contactless Heart Rate Measurement using Video Magnification,” accepted at 2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2022.

5. Better understanding of the effect of modern camera systems in smartphones -that are coupled with software features that apply changes to the videos post capture- on the Video Magnification algorithm.
6. Better understanding of the effect of two lighting conditions resulting in two illumination levels on the Video Magnification algorithm presented in an IEEE conference paper.

[11] L. Yassin Kassab, A. Law, B. Wallace, J. Larivière-Chartier, R. Goubran, and F. Knoefel, “Effects of Lighting and Window Length on Heart Rate Assessment through Video Magnification,” 2022, p. 6. In review for 2022 IEEE Sensors Applications Symposium Conference (SAS), 2022.

1.4 Thesis Organization

This thesis is organized in the following order: first, the current state of the literature of Video Magnification and its limitations are presented in Chapter 2 along with how this thesis will address some of these limitations. Second, the methods for addressing the limitations including video capture, ground-truth heart rate capture, and video analysis are presented in Chapter 3. After that, the results of the main contributions of the thesis are presented in Chapter 4 and Chapter 5 which are the effect of algorithmic parameters and the effect of noise on the Video Magnification algorithm, respectively. Finally, the thesis conclusion and future work are presented in Chapter 6.

Chapter 2: Background Review

This chapter first highlights the clinical significance of heart rate, it then provides a summary of recent advances and relevant works associated with the application of contactless heart rate measurement techniques over the last decade.

2.1 Heart Rate Signal from a Clinical Perspective

Heart rate is an important indicator of the health status of the heart organ which has two main tasks that run in parallel [1]. The first task is supplying oxygenated blood to the entire body from the lungs; allowing body organs to carry out their designated tasks [1]. The second task is pumping deoxygenated blood to the lungs from the entire body where the blood gets reoxygenated [1]. With every heartbeat, the heart performs these two mentioned tasks simultaneously, and the frequency at which the heart beats per minute is termed the heart rate.

Heart rate varies for different age groups, among different genders, and between rest and active states for the same human [1]. Factors that contribute to these variations include the size of the heart and the ratio of the size of the heart to the entire body [1]. The heart has muscles termed cardiac muscles that can expand by a process called hypertrophy; where cardiac cells expand in size as opposed to in number like other muscle types [1].

During physical exercise, the need for blood by muscles increases, requiring the heart to work harder to supply this higher demand of blood [1]. The heart does that by increasing its beating frequency and by expanding in size through hypertrophy [1]. Through these two modifications, a larger amount of blood can be supplied to the body organs that need it. Tracking heart rate changes over time can give an insight into the human's mental state as well as physical wellbeing [1]. Additionally, with age, monitoring heart rate becomes more important for early detection of illnesses and diseases including hypertension and diabetes [1].

2.2 State of the Literature on the Video Magnification Technique and its Limitations

Over the last decade, advances had been made in the area of contactless remote assessment of heart rate [4]. These advances began with the publication of two important works: Eulerian Video Magnification (EVM) [5] and Remote Photoplethysmography (rPPG) [6] in 2012 and 2013 respectively [4]. Both techniques are founded on the principle of detecting subtle colour changes in the skin captured by a video -which are imperceptible to the human eye- occurring as a result of the heart compression and recovery cycle [4].

Recent research has typically built upon one of the two techniques but not both, leading to parallel bodies of work that at first appear to define two distinct algorithms [4]. However, after further analysis, the two techniques appear to belong to the same family of video processing, and the distinction accounts more to the choices made at every step of the processing pipeline [4].

Eulerian Video Magnification (EVM) is a novel technique that enables the magnification of subtle colour and motion changes in videos that are imperceptible to the human eye [5]. This technique works by first spatially decomposing a video into different spatial frequency bands using a Laplacian Pyramid and then temporally filtering the pixel values using a custom-fit band-pass filter [5]. After that, the filtered signal is amplified using an amplification factor to reveal the minuscule colour and motion changes [5]. Using this technique, many researchers have shown it can non-invasively and unobtrusively detect heart rate in humans from their videos.

2.2.1 Summary of Review Papers on Remote Heart Rate Assessment

In the 2017 review [13], Hassan *et al.* summarized their review of the methods for heart rate measurement by dividing them into contact and non-contact methods (including methods based on EVM/rPPG techniques as well as other alternatives). Hassan *et al.* compared these methods according to their application, cost, robustness, immunity to noise, and technology maturation

statuses.

A review in 2018 specifically focused on the more challenging issue of vital sign assessment of the driver within an automotive environment [14]. Additionally, [14] covered many methods based on optical imaging techniques including visible light and thermal imaging to obtain those vital signs. Also in 2018 a review paper that compared a number of the variations of the rPPG algorithm was published [15] and although it is not a traditional review, it does provide a direct experimental comparison of a number of the rPPG methods that had been published to date using a publicly available dataset.

In 2019 a thorough review was published [16] that provided in-depth analysis of the current limitations of the non-contact vital sign detection techniques. The limitations include multiple regions of interest selection automatically for best results, multiple people monitoring simultaneously, long distance monitoring, and the use of multiple camera systems together. In 2019 a second review was also published [17] that was also extensive but focused generally on RGB (visible light) applications of the rPPG methods.

A third review in 2019 was published [18] that provided an extensive review of many vital signs over and above heart rate including blood pressure, respiratory rate, and oxygen saturation. It noted that in the papers working on the remote detection of vital signs, more research was done on detecting heart rate (75% of the papers), followed by respiration rate (39%) and then blood pressure (7%) and oxygen saturation (4%). This is a thorough review that considered single or multiple cameras and diverse camera types, although it was limited on fusion methods to combine different cameras.

The last review in 2019 [19] was a magazine style review that provided a high level summary of the various works on camera-based vital sign estimation from 2016-2018 with a tutorial style

background on the invention of the video-based monitoring methods. The two most recent reviews in 2020 are organized as deep technical tutorials on the algorithmic methods. The first review [20] built from the rPPG methodology with applications of RGB and IR cameras in contactless detection of vital signs. This review raised the issue of darker skin tones affecting performance and it referenced multiple databases of videos captured for analysis. The second review [21] on the other hand built from the EVM methodology and provided a review of the algorithmic methods as an entry point on EVM and compared the different methods at a high level.

In addition to these review articles, the PhD thesis documents for two students from Carleton university provide reviews of the literature generally focused on the EVM methodology using both RGB and thermal video techniques respectively [22], [23]. There has also been work in the literature that applies fusion-based methods of different cameras to estimate vital signs from facial video. These fusions include the combination of NIR with thermal camera [24], the combination of thermal imaging of the chest and face [25], and the combination of RGB, Depth and IR cameras for heart rate estimation [26].

2.2.2 Limitations of Contactless Heart Rate Measurement through Video

The Video Magnification technique faces many challenges including its sensitivity to motion [27], [28], and light conditions [16], [17], [29], and its bias to skin tone [16], [28] that all negatively impact performance. The Video Magnification technique used in this thesis is an implementation based on EVM and it will be described in detail in Chapter 3 under Video Analysis. Many different camera types have been examined with EVM for the purpose of heart rate extraction with different advantages and challenges of each.

These camera types include optical visual light: RGB and Monochrome cameras, [6], [30]–[35], thermal cameras [29], [36]–[40] that use mid to long wavelength infrared (IR), and IR cameras

[41] that use short wavelength IR. Optical cameras are sensitive to light conditions and skin colour. Thermal cameras have the advantage of working in the absence of light, providing temperature information, and identifying facial features such as vasculature. IR cameras are less expensive than thermal cameras, work in the absence of light, and are less dependent on skin tones. For the skin tone limitation, one work in the literature introduced a skin reflection model to manage skin and light variation [42] while another work analyzed image background to estimate and compensate for lighting shift [43].

There are many possible regions of interest (ROI) for heart rate detection with EVM including the forehead and cheeks on the face [31], the neck, and the wrist [30]. However, [38] showed that mean intensity values in a ROI on the forehead better correlated with the ECG signal than an ROI on the wrist. One work used face detection and tracking to obtain and track a ROI [44] while another work tried the fusion of RGB and thermal cameras by applying face tracking to the RGB camera and using that ROI for the thermal camera analysis [45]. Nevertheless, no work has identified the best ROI for heart rate detection and therefore the selection of the ROI, which is the region of skin analyzed by the video magnification algorithm is an area that requires further exploration. The choice of ROI can be impacted by issues such as hair occlusion on the face, face covering with a mask, or head motion.

Another area that requires investigating is the size of the time window segment that is analyzed for heart rate extraction. Different window segment lengths had been used in the literature including 4 seconds [46], 10 seconds [47] and 30 seconds [35], [48]. However, no work has investigated the effect of varying the window segment length on the Video Magnification algorithm nor identified the best window segment length to be used for the analysis.

Noise is another limitation of the Video Magnification algorithm, and it comes from different

sources including subject motion, thermal noise, quantization noise, and noise from changes in light illuminations as a result of subject motion. Work has been done in the literature to address subject motion challenges including pre-processing [28] and post processing [49] analysis applied to videos. However, no work has previously quantified the limitation of the Video Magnification algorithm in the presence of noise.

Additionally, different lighting sources have been used with visible colour cameras including ambient light [34], [36], [50], incandescent and fluorescent light [29] and LED light [51]. However, there is no work that determined the best lighting source for contactless heart rate monitoring nor compared and evaluated the spectral properties of different light sources. Additionally, there is a lack of focus on the effect of combining multiple light sources together.

Lastly, an area that has not been investigated in the literature is the use of modern smartphone camera systems with Video Magnification to extract heart rate. Modern smartphone camera systems have advanced lenses that enable higher quality video capture. However, these videos are being processed by GPUs with automated algorithms that modify the pixel intensities in the videos post capture to enhance the video appearance. The effect of these software enhancements on the videos have not been investigated for contactless heart rate detection using the Video Magnification technique.

2.3 Work to Address the Limitations of Video Magnification

In this thesis, the effect of the region of interest (ROI) selection on Video Magnification is first investigated, followed by the window segment length. After that, the effect of noise on Video Magnification is investigated followed by the effect of modern smartphone camera systems and lastly the effect of the light source is investigated.

In the next chapter, the methods for recording the videos used to generate the results of this thesis

along with the methods for ground-truth heart rate measurement, and the methods for analyzing the videos are detailed.

Chapter 3: Method and Experimental Setup

This chapter first details the methods for three video capture processes used to generate the results of this thesis. It then details the methods for the ground truth heart rate measurement that the results were compared to. Finally, it details the video analysis method used to process all videos.

3.1 Video Capture:

The videos used to generate the results of this thesis came from one of three sources. The first source is from a collaboration with the National Research Council's Aerospace Research team under the Aerospace Portfolio through the Pandemic Response Challenge program. The second source is from self-captured videos at home of myself. And the last source is artificially generated videos in MATLAB.

3.1.1 Video Capture Method from the National Research Council Collaboration.

As part of the collaboration with the National Research Council's Aerospace Research team under the Aerospace Portfolio through the Pandemic Response Challenge program, NRC has setup and run an experiment collecting videos of 22 subjects under 14 different conditions. This was done as part of NRC's Multi-Modal Contactless Vital Sign Assessment Solution for potential use at airports [10], [11], [52]–[54] that had an ethics protocol with REB # 2020-150 [54] .

Three imaging modalities of RGB colour, thermal, and monochrome cameras were used for the data collection. Nine cameras in total were setup in the anechoic chamber at the National Research Council that included:

- RGB Cameras
 - Network security RGB Panasonic WV-S1131 camera
 - Basler aca1300-60gc RGB camera
- Thermal Cameras

- FLIR SC8303 MWIR thermal camera
- FLIR T650sc LWIR thermal camera
- Teledyne DALSA caliber LWIR thermal camera
- Monochrome Cameras
 - Four Basler aca1300-60gm monochrome cameras at wavelengths 575nm, 675nm, 750nm, and 850nm.

However, only the RGB Panasonic WV-S1131 camera videos were used to generate the contributions of this thesis. 22 subjects were recruited for this research experiment. The recruited subjects had varied age (21-73 years old), skin tones (very fair (3), fair (7), olive (1), medium (6), dark (3), and black (2)) and included both genders (4 females, 18 males). Two lighting sources were used: LED and Halogen. LED lighting was provided by two LED panels mounted on an aluminum frame and located 2 meters in front of the participants at a height of 1.9 meter. Halogen lighting was provided by two 300W tungsten-Halogen bulbs placed inside Silverbox Quartz Softboxes and the front of the Softboxes was overlaid with the Lee 250 half white diffusion filters. Each of the two Halogen lights was placed directly under the LED light panel. The experimental conditions varied three parameters: the subject's pose between sitting and standing, the lighting conditions between LED only and LED and Halogen together, and whether the subject had a mask on or off. A Summary of 8 out of the 14 conditions which are used in this thesis for vital signs assessment is in Table 1 on the next page. The captured data included 90-second videos of each subject for each experimental condition using all nine imaging cameras. Ground truth heart rate readings were recorded using a PPG device worn by the subjects as discussed in section 3.2 below.

Table 1. Summary of the experimental conditions for the tests done by NRC

Condition Label	Position (Sitting/Standing)	Lighting (LED/LED+Halogen)	Mask (On/Off)
C1	Sitting	LED	Off
C2		LED+Halogen	
C3		LED+Halogen	On
C4		LED	
C5	Standing	LED	Off
C6		LED+Halogen	
C7		LED+Halogen	On
C8		LED	

The experimental conditions were chosen to simulate real world scenarios at airports or other crowd-gathering places where people would be either sitting or standing. Additionally, with the onset of Covid-19, the need for wearing masks emerged and this was included in the experiment design for research purposes. The ability to monitor heart rate using remote non-contact methods would be advantageous in crowd-gathering places because it would be unobtrusive to individuals, and it would decrease the need for interaction with trained personnel.

3.1.2 Self-Captured Videos Method

The other source of videos used to generate the results of this thesis were self-captured videos of myself (25 years old, fair skin, female). Self-captured videos were recorded using an iPhone 12 Pro Max camera system and a VGA web camera. This iPhone camera system is made up of three cameras: Ultra-Wide, Wide and Telephoto summarized in Table 2 with their aperture information. The smart software features of the iPhone 12 Pro Max are summarized in Table 3. The features include a High-Definition Recording (HDR) video option, an auto 30FPS option, an auto 30&60 FPS option, an option to switch off the default setting of switching between the three cameras

during video capture, and an option that combines HDR mode with auto 30&60. Additionally, all these five features could be turned on and off manually at the same time.

Table 2. Smart camera system iPhone 12 Pro Max

Camera Type	Feature Description
Ultra-Wide	<i>f/2.4</i> aperture and 120° field of view
Wide	<i>f/1.6</i> aperture
Telephoto	<i>f/2.2</i> aperture

Six different videos were recorded subsequently of myself for 30 seconds each while avoiding any unnecessary motion. The first video was taken with a Logitech C270 VGA Webcam, that records HD videos at 720p and 30FPS. The subsequent five videos were taken using an iPhone 12 Pro Max placed in the same position as the VGA camera and set to record videos at 720p and 30FPS. The lighting was provided by natural light from a large window.

For the iPhone videos, first all smart features in Table 3 were manually turned off and the camera lock feature was turned on so the recorded video would not contain frames from different cameras combined. Next, each of the four remaining smart features in Table 3 were turned on separately and a video was recorded.

The captured videos were then processed using the Video Magnification algorithm with input parameters set to process the 30 seconds of video: 50-150 BPM bandpass parameters, 10 second window segment length with 90% overlap and FFT size of 2048. The results were then compared to the ground truth heart rate to evaluate performance. The region of interest was of size 64x64 pixel centered on the forehead.

Table 3. Summary of the smart video features in iPhone 12 Pro Max

Feature	Feature Description
HDR Video	Records up to 60 FPS video in 10-bit high dynamic range including Dolby Vision
Auto 30 FPS	Automatically reduces the framerate to 24 FPS to improve low light video and to optimize the size
Auto 30&60 FPS	Automatically optimizes the framerate of 60 FPS video based on scene motion
No Camera Lock	Automatically switches between cameras while recording video
HDR + Auto 30&60	Combines the two features together

3.1.3 Artificially Generated Videos in MATLAB

The last source of videos used to generate the results of this thesis were artificially generated video in MATLAB. The videos were generated using the MATLAB Image Processing Toolbox. The video generation process could be further divided into two sections, the first one does not include noise, while the second one includes White Gaussian Noise.

3.1.3.1 Video Generation without Noise

The generated videos were 10 seconds in duration with a framerate of 30 FPS. They contained a sine signal that modifies the green colour pixel intensities only based on Equation 1. In the equation, k is the signal amplitude, HR is the heart rate frequency to be simulated, and n ranges from 1 to 301 (to span the total number of frames in the video $10s * 30 FPS = 300$ frames). The equation contains a DC offset of 0.5 to place the signal within the range of the integer values in the video as the pixel values are between 0 and 1 prior to conversion to unsigned integers.

Equation 1. Simulated sine signal equation for artificial video

$$\text{simSig}[n] = 0.5 + k \sin\left(\frac{2\pi HRn}{60}\right)$$

Four heart rate frequencies were simulated: 53, 60, 73, and 83 BPM. They were chosen as normal

heart rate frequencies that are not harmonics of each other or the video framerate. Each frame of the video was a 200x200 pixels image that contained one of the green colour intensities from Equation 1. The single 300 images were then converted into video frames by MATLAB and combined to create the video. Figure 1 visually illustrates the video generation process.

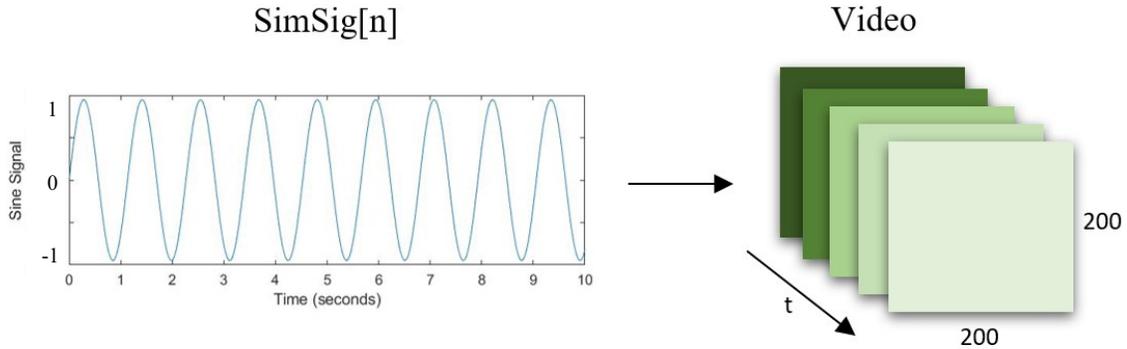


Figure 1. Illustration of the video generation method

Parameter k is what was varied to distinguish large and small signal amplitudes. For the large signal amplitudes, k ranged from 0.01 to 0.05 in a 0.01 step size. This range was denoted large because the change in the green colour pixel intensities was visible to the naked eye in the videos.

For the small signal amplitudes, k ranged from 0.002 to 0.005 in a step size of 0.001. This range was denoted small because the change in the green colour pixel intensities was invisible to the naked eye in the video. This range covers the desired signal amplitudes for investigation with the VM algorithm performance as it is representative of the unperceivable subject heart rate signal analyzed in real video. The standard representation of values in videos is 8-bit integers, therefore when the signal SimSig in Equation 1 is converted to integers, the value of k is inter-related with the quantization levels. Since the signal range is 0 to 1 prior to the conversion, then the quantization spacing is $1/(2^8)$, which equals 0.00391. Therefore, signal levels between ± 0.00196 ($0.00391/2$) within a quantization level would be rounded to that level. A signal amplitude $k = 0.002$ would fall

just above one quantization level.

3.1.3.2 Video Generation with Noise

Additionally, videos containing White Gaussian Noise were generated according to Equation 2 of 10 second duration. The videos were generated in the same method described in the previous subsection, but a value from a quantified noise level was added at each signal point. MATLAB's White Gaussian noise function "wgn" was used to generate a matrix of the specified size 300x1 that generates White noise at a power level 0.005 given by Equation 3 . This power level was chosen as half of the smallest large signal amplitude, which was 0.01. The power level was scaled by factor m . A visual illustration of the video generation from the signal $\text{simSigN}[n,m]$ can be seen in Figure 2.

Equation 2. Simulated sine signal equation for artificial video with noise

$$\text{simSigN}[n, m] = \text{simSig}[n] + \text{simNoise}[m]$$

Equation 3. Simulated sine signal equation for artificial video with noise

$$dB_{wgn}(m) = 10 \log_{10}(0.005 * m)^2$$

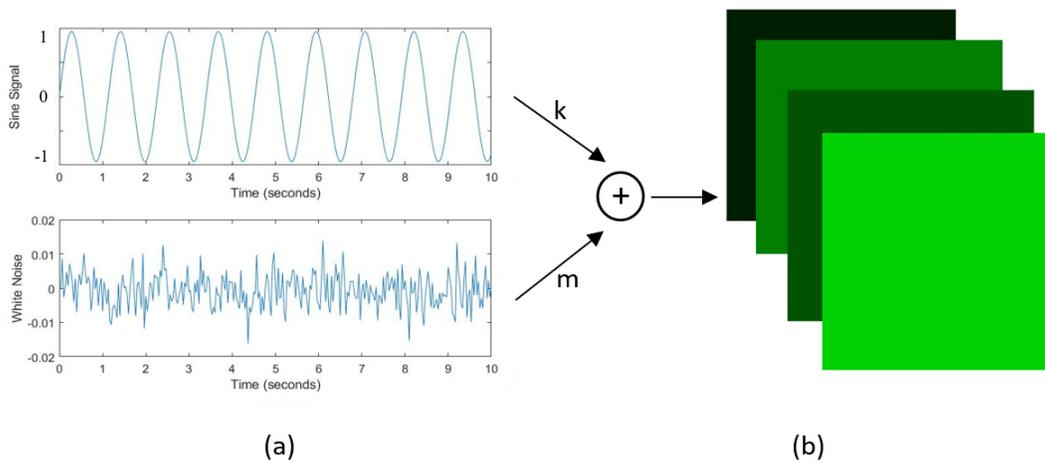


Figure 2. Illustration of the video generation method with noise obtained from [12]

The values of m ranged from 5 to 30 in a step size of 5 for a total of 6 noise levels. These noise levels correspond to a noise power range of 6.4 to 38.4 quantization levels.

3.2 Ground-truth Heart Rate Capture

The ground truth heart rate readings used to compare the results of the Video Magnification algorithm were captured using a pulse oximeter for the NRC group videos and the self-captured videos. The details are further provided in this section.

3.2.1 Ground-truth for the NRC Video Capture

True heart rate readings were collected from the subjects using a SenSmart pulse oximeter that collected photoplethysmography (PPG) signal. The pulse oximeter interfaced to the data acquisition system through a Nonin 3500SP signal processor and SenSmart Model X-100H hub. The system times of the acquisition computer and the cameras were synchronized using a network time protocol server.

3.2.2 Ground-truth for the Self-Captured Videos

For the self-captured videos, true heart rate readings were collected using a fingertip pulse oximeter blood oxygen saturation monitor by the company HealthTree that interfaces to the phone through an app. Results were provided graphically and numerically for comparison.

3.3 Video Analysis Method

The Video Magnification technique used to process all videos and generate the results of this thesis is detailed in this section.

Videos were processed using a Video Magnification algorithm that had been implemented based on the EVM technique by previous graduate students working in the same lab at Carleton university [27], [38], [55]. This algorithm was taken and turned into an app in Python by the lab technician. This app has a user interface for the manual processing of a video in addition to a programmatic option for the automatic processing of many videos. The Video Magnification algorithm computes the heart rate of a subject using a video of the subject. The output of the VM algorithm includes four plots containing a time-series signal, an estimated heart rate over time, a spectrogram containing all the frequency components of the time-series signal and a plot containing computed FFTs for all window segments in the video. The VM algorithm takes in multiple parameters in addition to the video as an input. These parameters are entered manually, and they include: the video start and end times, the cutoff frequencies for the band pass filter (lowest and highest heart rate frequencies to include), the window segment length, the segment update time (the # of seconds to skip between adjacent windows), the FFT size, and the location of the region of interest (ROI). The ROI is manually selected on the first frame of the video using a cursor and it is used for every subsequent frame in the video afterwards. Figure 3 shows the different inputs and outputs of the VM algorithm. Varying any of the input parameters influences the performance of the VM

algorithm as discussed later Chapter 4.

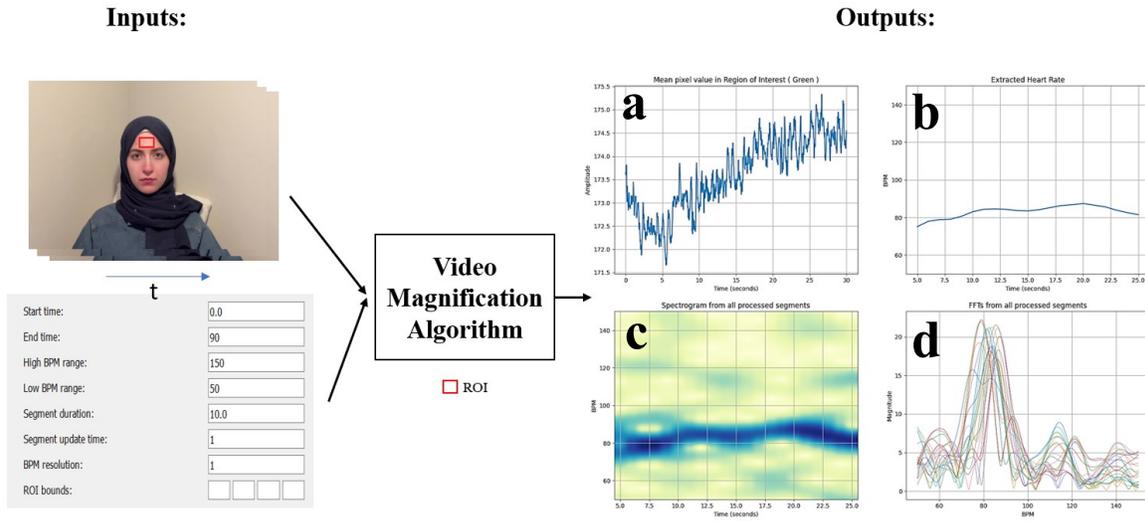


Figure 3. Inputs and outputs of the Video Magnification algorithm using a video frame of the author

Taking a video recorded at “x” frames per second (FPS) for a duration of “y” seconds, the video would contain $x*y$ frames. To measure the heart rate in the video, the video is divided into smaller window segments of size “z” seconds with frames overlapping between adjacent windows. If the video is recorded at 30 FPS for 90 seconds and the window segment length is set to 10 seconds with 90% overlap between adjacent window segments, then the total number of video frames would be 2700, the total number of windows would be 81, and each window would contain 300 frames. Next, three time-series signals are obtained by computing the average of the red/green/blue channels pixel intensities separately contained within the ROI for each video frame in the window segment. Next, for each window segment, a Hamming window is applied to the signal and an FFT is then computed for this window segment. Following this step, the FFT gets reduced according to the cutoff frequencies of the band pass filter. The final step is computing the average heart rate for that window which corresponds to the peak value in the computed FFT for that window. The same process is repeated for every window segment in the video. For each window, an average is taken

of the computed average pixel intensities for each frame in the window. This average value taken from each window results in the time-series signal for the entire video presented in a plot as an output of the VM algorithm. Figure 4 summarizing the various steps of the Video Magnification algorithm involved in the analysis.

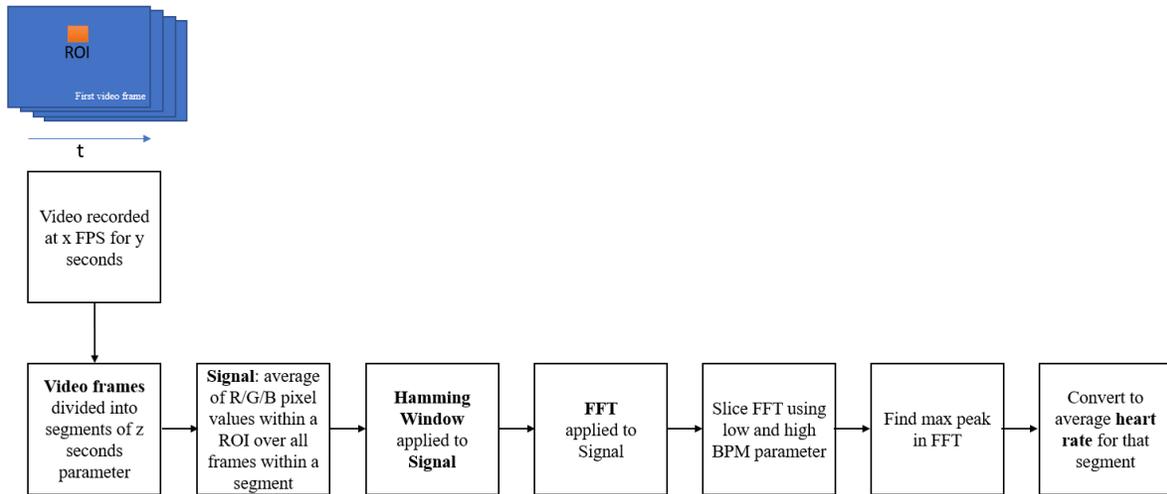


Figure 4. Flowchart containing the detailed steps of the Video Magnification algorithm

As shown in Figure 3, the output of the Video Magnification algorithm contains 4 different figures. The first figure (a) contains the mean pixel value in the region of interest for the selected colour channel over the whole video duration. Another figure (d) contains the computed FFTs for all the window segments in the video, while a third figure (b) contains the extracted heart rate that corresponds to the peak of each window segment FFT. Finally, the last figure (c) is a spectrogram that shows all detected frequencies within the region of interest.

In the next chapter, the effects of varying two of the mentioned input parameters: the ROI size and the window segment length are detailed.

Chapter 4: Effects of Algorithm Parameters

In this chapter, first the effect of varying the region of interest size on the performance of the Video Magnification algorithm is detailed. After that, the effect of varying the window segment length on the performance of the Video Magnification algorithm is detailed.

4.1 Effect of Region of Interest Size on the Video Magnification Algorithm.

This section's contents are based on the paper published at the 16th edition IEEE International Symposium on Medical Measurements and Applications in 2021 titled *Effects of Region of Interest Size on Heart Rate Assessment through Video Magnification* [10].

The region of interest (ROI) is the region of skin chosen by the user that is analyzed by the Video Magnification algorithm as described earlier. The choice of ROI can be impacted by factors such as hair occlusion, face masks and head motion when the user face is not tracked. All three factors can lead to regions in the face being constantly visible while other regions moving out of the ROI frame. Previous research has considered small size ROIs or ROI subset selection for heart rate estimation [16], [27]. The work described in this section however details the heart rate measurement accuracy of Video Magnification when using very small ROIs as well as large ROIs.

4.1.1 Analysis Details

In order to determine the effect of the ROI size on the accuracy of the VM algorithm in detecting heart rate, four videos obtained from the NRC dataset described earlier corresponding to conditions C1, C4, C5 and C8 from Table 1 were used for 1 subject. The subject was one of the co-authors in [10]. These videos were captured by the RGB Panasonic WV-S1131 camera. These conditions varied 2 parameters: pose (sitting vs. standing) and face covering by a mask (mask on vs. off). The source of lighting for the four videos was Light Emitting Diode (LED) light provided by two LED panels. Table 4 summarizes the four conditions again.

Table 4. A summary of the four testing conditions used in the analysis [10]

Condition:	Sitting or Standing	Mask On or Off
C1	Sitting	Off
C4	Sitting	On
C5	Standing	Off
C8	Standing	On

For conditions C1 and C4, the subject was sitting on a chair in front of the camera and for conditions C5 and C8, the subject was standing in front of the camera. In condition C4 and C8, the subject was wearing a face mask that covered his nose, mouth and parts of the cheeks. In all four cases, the subject was asked to minimize movement as much as possible during a 90-second video capture time for each condition.

The four videos were run through the VM algorithm with fixed input parameters for video length (90s), band pass cut-off frequencies (50BPM to 150BPM), window segment length (10s) and 90% overlap between adjacent windows. The only varied parameter in the analysis was the ROI, which was square in shape. The chosen ROI location was fixed across all video frames, and it was selected on the first frame of the video.

The size of the ROI was varied to include very small sizes and large sizes. Small ROI sizes included 1x1, 2x2, 4x4, 8x8, and 16x16 pixels and they were all located on the forehead region. Large ROI sizes included 16x16, 32x32, 64x64 (all on the forehead), 128x128 (covered the forehead and eyes), 256x256 (covered the full face), and 512x512 (covered the full face, neck and included background). The different areas covered by the large ROI sizes can be seen in Figure 5 obtained from [10].



Figure 5. Areas covered by the different large ROI sizes: top left: 16x16, bottom left: 32x32, top right: 64x64, bottom right: 256x256, middle center: 512x512 from [10]

For the small ROIs, the average of 16 contiguous blocks of pixels for each ROI size was considered for the analysis. Therefore, the analysis for small ROI sizes 4x4, 8x8 and 16x16 covered the same regions included in the large ROI sizes 16x16, 32x32 and 64x64 respectively. A visualization of the mapping between small and large ROIs can be found in Figure 6. This is interesting because in instances when there is no single large ROI to analyze, being able to combine small regions from around the face into a larger ROI may be helpful.

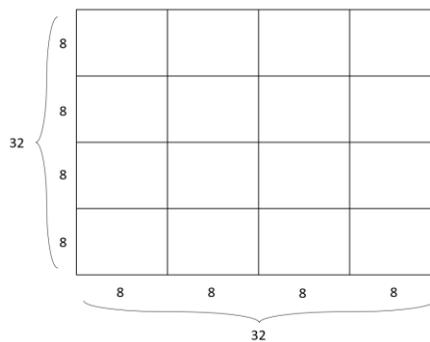


Figure 6. Relationship between the small ROI and the large ROI analysis (small 8x8, large 32x32)

In the 8x8 small ROI case, the estimated heart rate from the 16 contiguous 8x8 squares were separately evaluated and then averaged together. These 16 contiguous squares cover the same

region contained in the large ROI size of 32x32. Therefore, in total, each of the four videos was processed 10 times to obtain the heart rate of the subject. The results were then compared to the subject's true heart rate obtained from the pulse oximeter described earlier. The percent mean absolute error metric according to Equation 4 was used to compare the HR results obtained from the VM algorithm to the true HR from the pulse oximeter. The reason for that is that the algorithmic error could be either positive or negative and the percent mean absolute error would consider all errors without averaging across them.

Equation 4. Percent mean absolute error metric for heart rate

$$\% \text{ Mean Abs Error} = \left| \frac{\text{Video Mag HR} - \text{True PPG HR}}{\text{True PPG HR}} \right| \times 100$$

4.1.2 Effect of Small Region of Interest Size:

Figure 7, Figure 8, and Figure 9 show the percent mean absolute error for the small ROI sizes corresponding to the green, red, and blue colour channels respectively.

Out of the three colour channels, the best performance corresponds to the green colour and previous work had identified the green channel as having the better performance in heart rate detection from video over the other two colours [56], [57]. For all three channels, the sitting cases (C1 and C4) yielded better results with a smaller % mean absolute error than the standing cases (C5 and C8). The degradation in performance for the standing cases is likely related to movement artifacts. Additionally, the error for the sitting cases C1 and C4 although expected to be the same since the ROIs are on the forehead and do not include the face mask, was lower in C1 (mask off) than C4 (mask on).

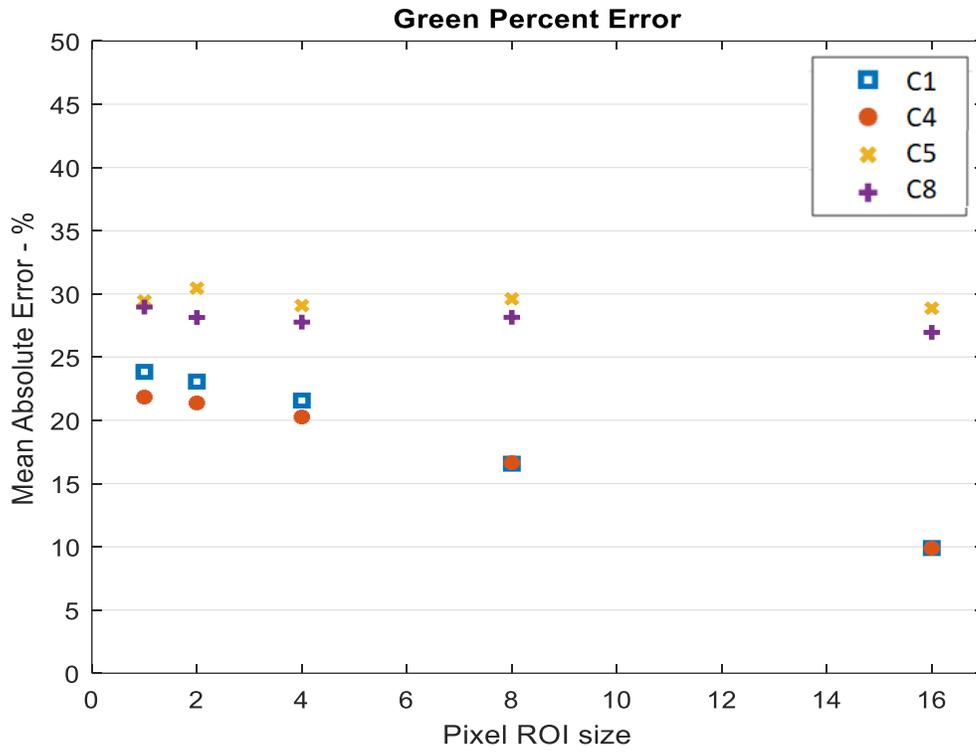


Figure 7. % mean absolute error for average of the 16 contiguous small ROIs- green from [10]

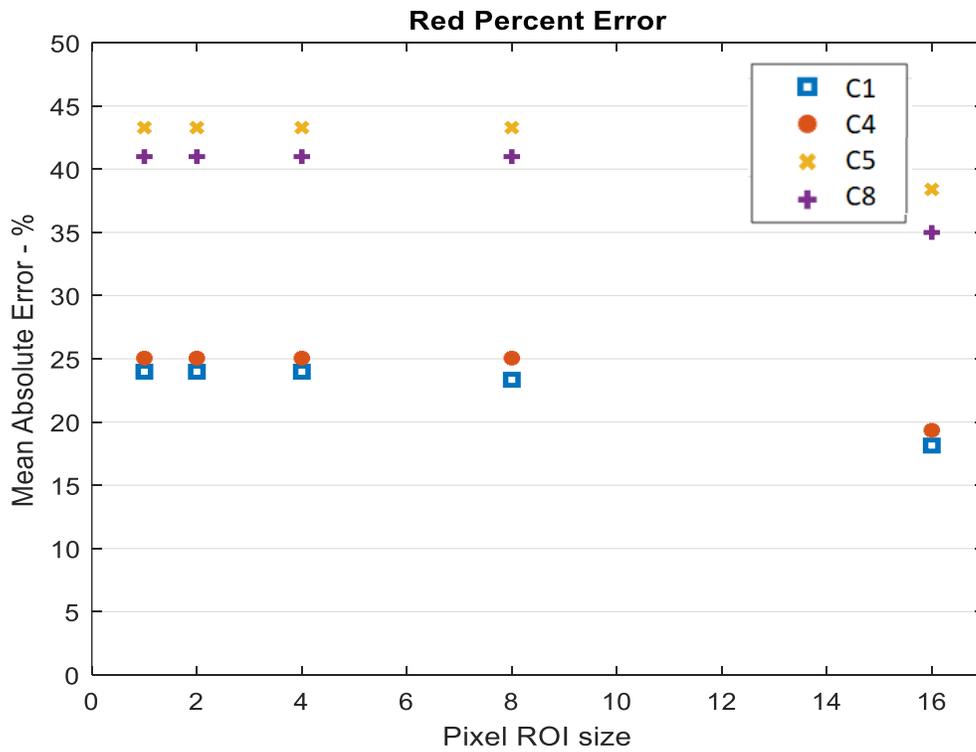


Figure 8. % mean absolute error for average of the 16 contiguous small ROIs - red from [10]

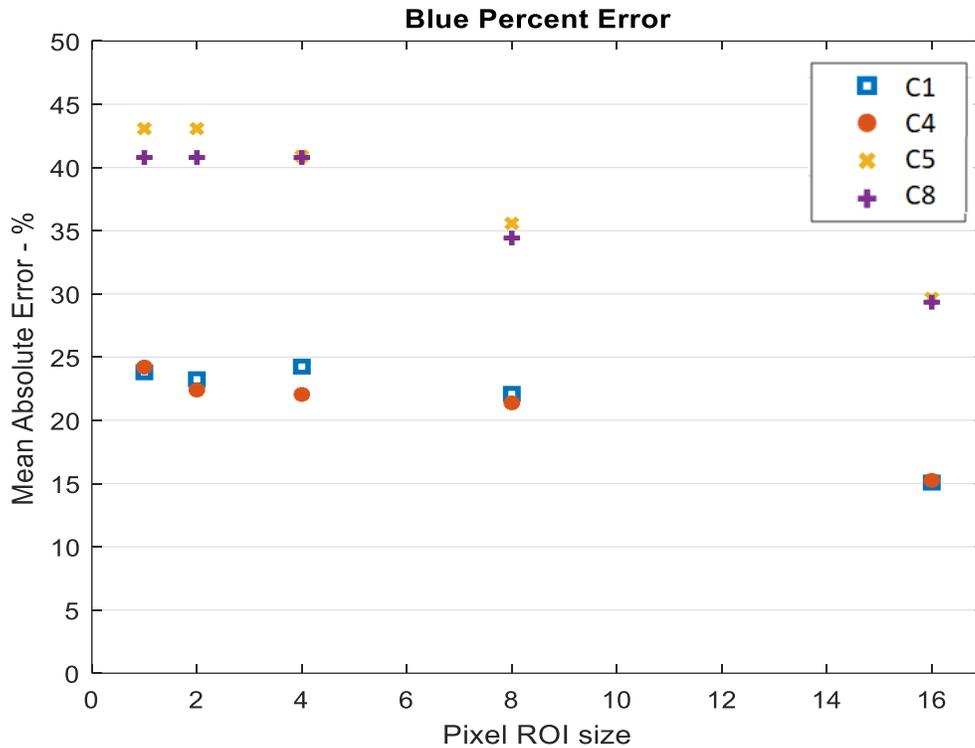


Figure 9. % mean absolute error for average of the 16 contiguous small ROIs - blue from [10]

The worst results for all four conditions correspond to the red channel. The blue channel results lie in the middle between red and green where it is the same as red for the standing cases (C5 and C8) and slightly better than red for the sitting cases (C1 and C4). It is important to note as well that as the ROI size increases, the performance of the VM algorithm in detecting heart rate improves for all three colour channels. Additionally, the increase in motion between the sitting and standing cases has a detrimental effect on the accuracy of the VM algorithm in detecting heart rate for all colour channels and all ROI sizes.

4.1.3 Effect of Large Region of Interest Size

Figure 10, Figure 11, and Figure 12 show the percent mean absolute error for the large ROI sizes corresponding to the green, red and blue colour channels respectively.

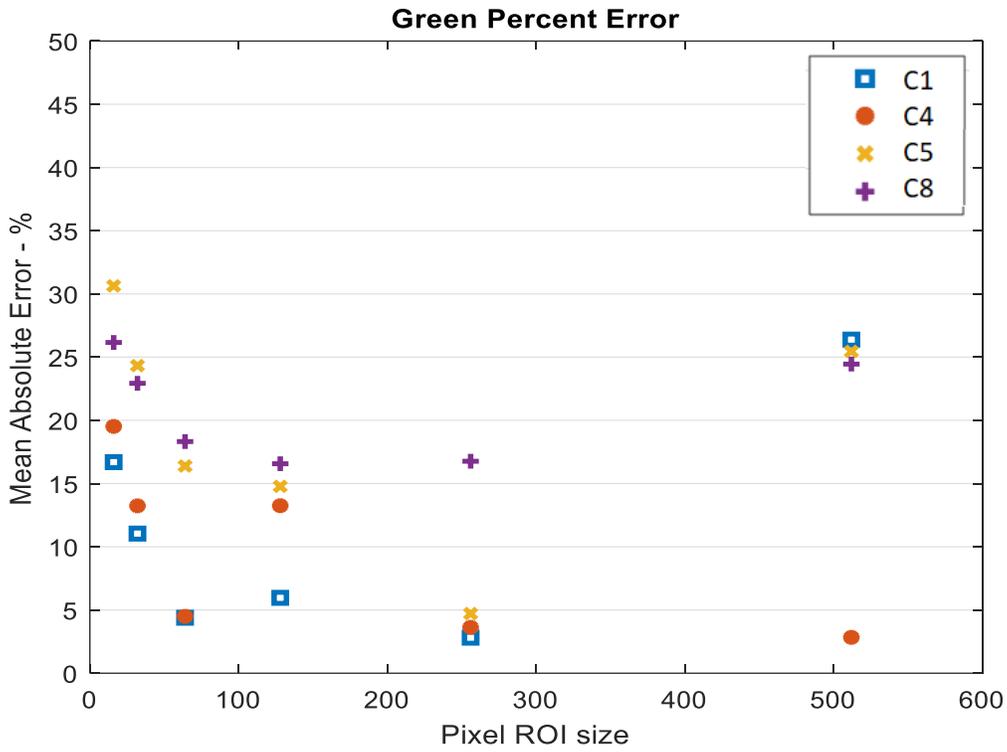


Figure 10. % mean absolute error for large ROIs - green channel from [10]

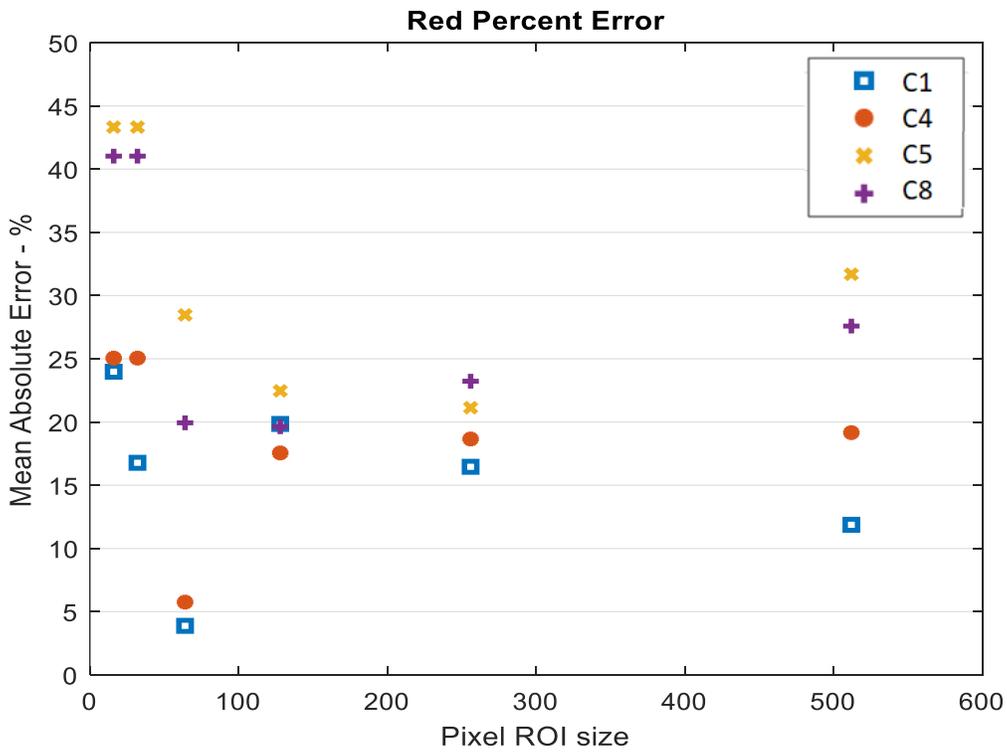


Figure 11. % mean absolute error for large ROIs - red channel from [10]

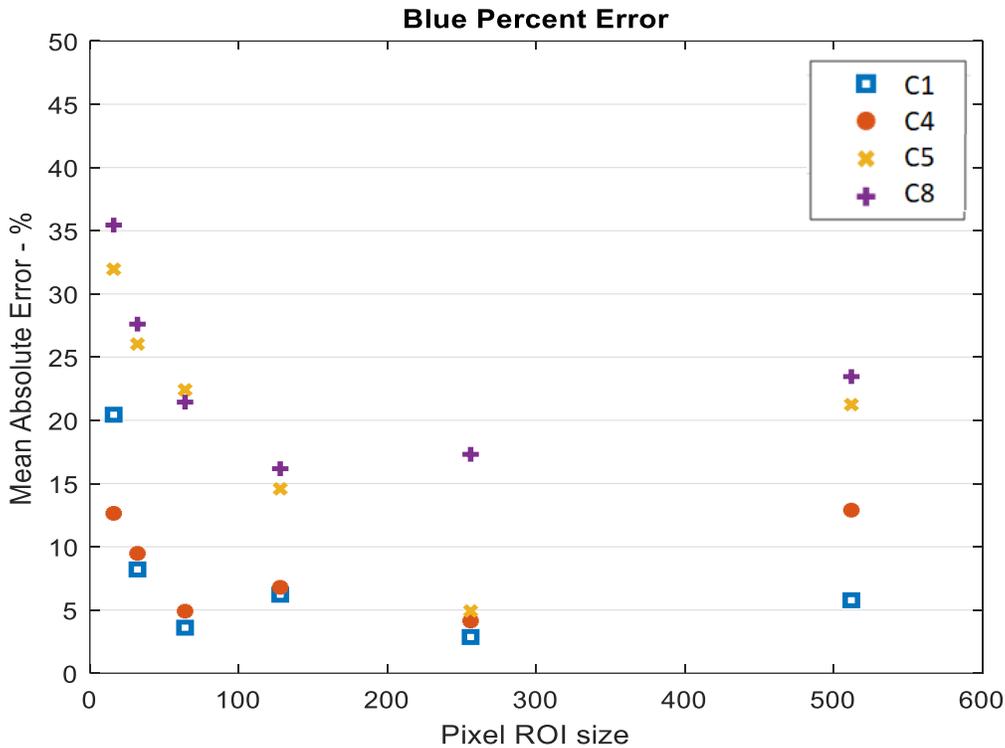


Figure 12. % mean absolute error for large ROIs - blue channel from [10]

Similar to the case with the small ROIs, the green channel results are the best out of the three colour channels for all four conditions. The red channel results are the worst, and the blue channel results fall in between green and red. Additionally, the blue channel results approach the green channel results in many cases. The results for the sitting cases (C1 and C4) for all three colours are better than the standing cases (C5 and C8) as a result of motion.

For the large ROI sizes of 16x16, 32x32, and 64x64 (all within the forehead), the green channel results respectively improved as the ROI size increased for all four cases as shown in Figure 10. For the 128x128 ROI size that contained the eyes in addition to the forehead, although it was bigger than 64x64, the green channel results degraded. This degradation in performance is attributed to eye blinking presenting as a motion artifact.

The green channel results improved again for the bigger ROI size of 256x256 that encompassed the entire face, thus adding cheeks and parts of the neck containing a pulse to the analysis. The

results for the large ROIs (16x16-128x128) are expected. However, the green 256x256 results for the sitting cases (C1 and C4) were very similar which indicates that the face mask did not affect the results.

For the same ROI (256x256), condition C8 (standing with mask on) had the highest percent mean error. This highlights the effect of combining motion and a face mask on the degradation of the accuracy of heart rate measurement using Video Magnification. Lastly, the largest ROI size (512x512) that included the background in the analysis, resulted in a significant decline in the accuracy of the heart rate measurement from video. The effect of subject motion would be most obvious in this large ROI size, and combined with the background, the two explain the decline in results.

4.1.4 Comparing Small and Large ROI Sizes

As previously mentioned, the small ROIs and large ROIs were processed in two different ways. The heart rate computed from the small ROIs was the average of the results from 16 contiguous square blocks of pixels corresponding to the ROI size. While the heart rate computed from the large ROIs was computed once from a square block the size of the ROI. Therefore, the analysis for small ROI sizes 4x4, 8x8 and 16x16 covered the same regions included in the large ROI sizes 16x16, 32x32 and 64x64 respectively. When the results from the two methods are compared, the large ROI method outperformed the small ROI method for all cases. This indicates that the average of all pixel values within a ROI should be compared before estimating the heart rate rather than estimating the heart rate separately on different adjacent ROIs.

4.1.5 Continuous Absolute Error for the Four Conditions

The green colour channel analysis results at a ROI size of 256x256 are further detailed and discussed in this section for the four test conditions C1, C4, C5 and C8. Figure 13 shows the

continuous absolute error over the four 90 seconds videos for each 10 seconds window analyzed. In this figure, it can be seen that C1 (sitting with no mask) has the smallest deviation from the true heart rate for the subject, followed by C4 (sitting with mask), which shows larger deviations. The errors for C5 and C8 (standing cases) are much higher, highlighting the effect of motion in decreasing the performance of the VM algorithm.

The VM algorithm results (quad image described in the Video Analysis Method section in Chapter 3) for C1 and C4 can be seen in Figure 14 and Figure 15 respectively, while the results for C5 and C8 can be seen in Figure 16 and Figure 17 respectively. These four graphs show the effect of motion (standing cases) in decreasing the VM algorithm performance.

In Figure 14 and Figure 15 the algorithm's ability at accurately detecting the subject's heart rate of 60BPM with a very small deviation is highlighted. On the other hand, in Figure 16 and Figure 17 it can be seen that the algorithm detects the true heart rate of 80BPM sometimes, but it does not detect it other times. It is interesting to note that the spectrograms of Figure 16 and Figure 17 contain a continuous heart rate of 80BPM but they also contain other frequency components that are likely due to head and body motion, causing the algorithm's detected heart rate to deviate from the true HR.

Lastly, the absolute error for C8 in Fig 16 shows positive as well as negative errors that could cancel each other when averaged, resulting in a smaller average percent mean absolute error for that condition over C5.

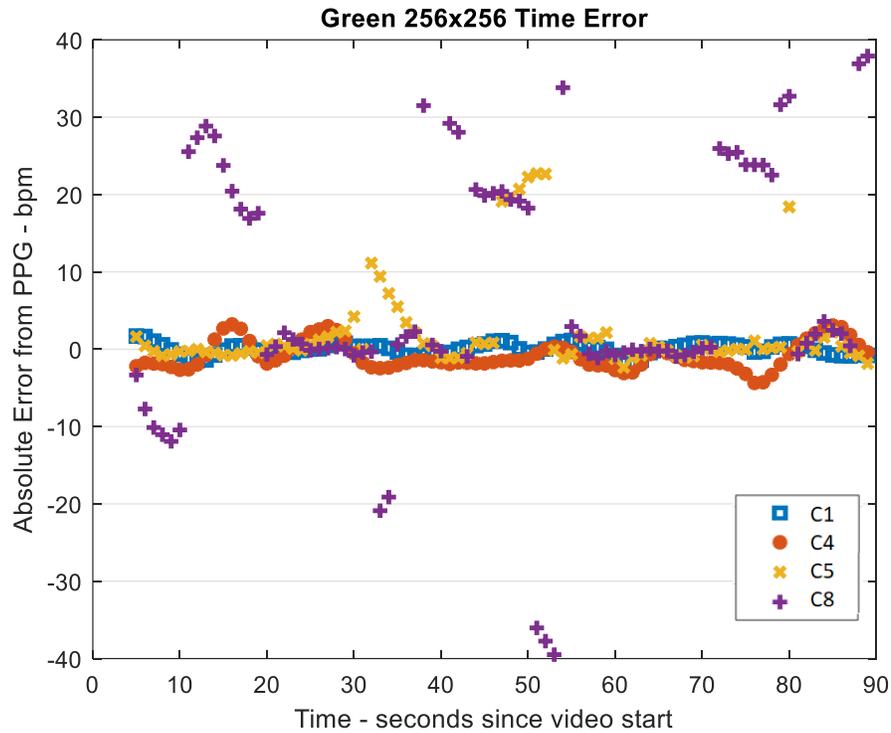


Figure 13. Continuous % mean absolute error of C1,C4,C5,C8, green channel, ROI 256x256 from [10]

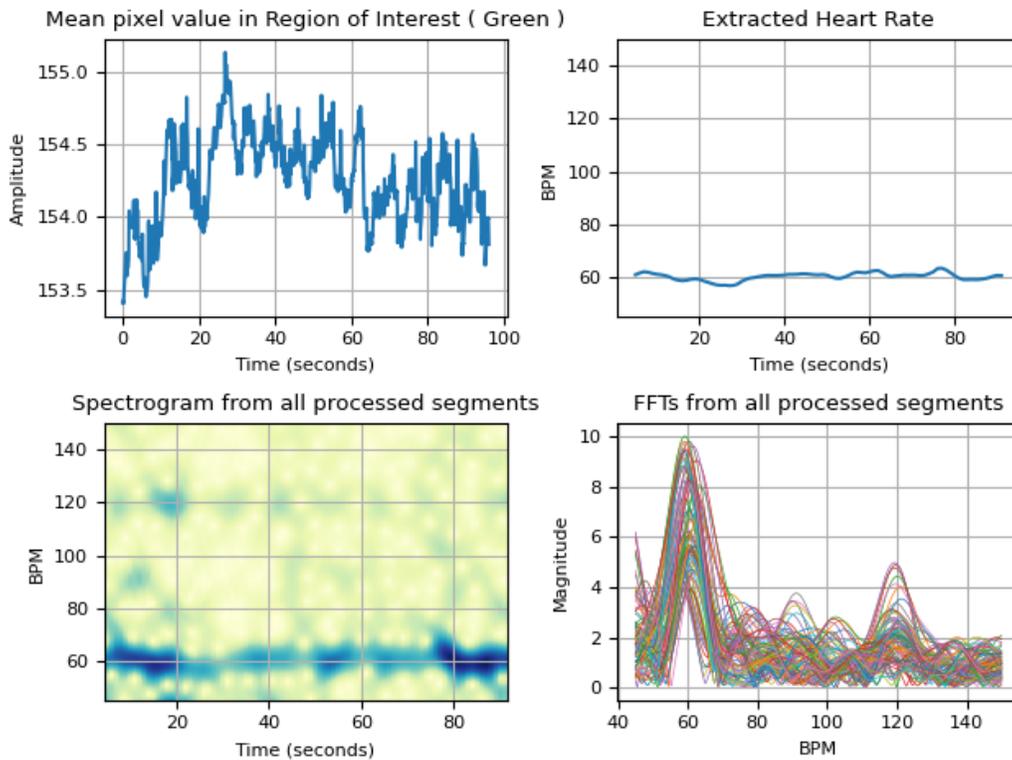


Figure 14. VM algorithm results for condition C1 (sitting, no mask), green, 256x256 ROI, from [10]

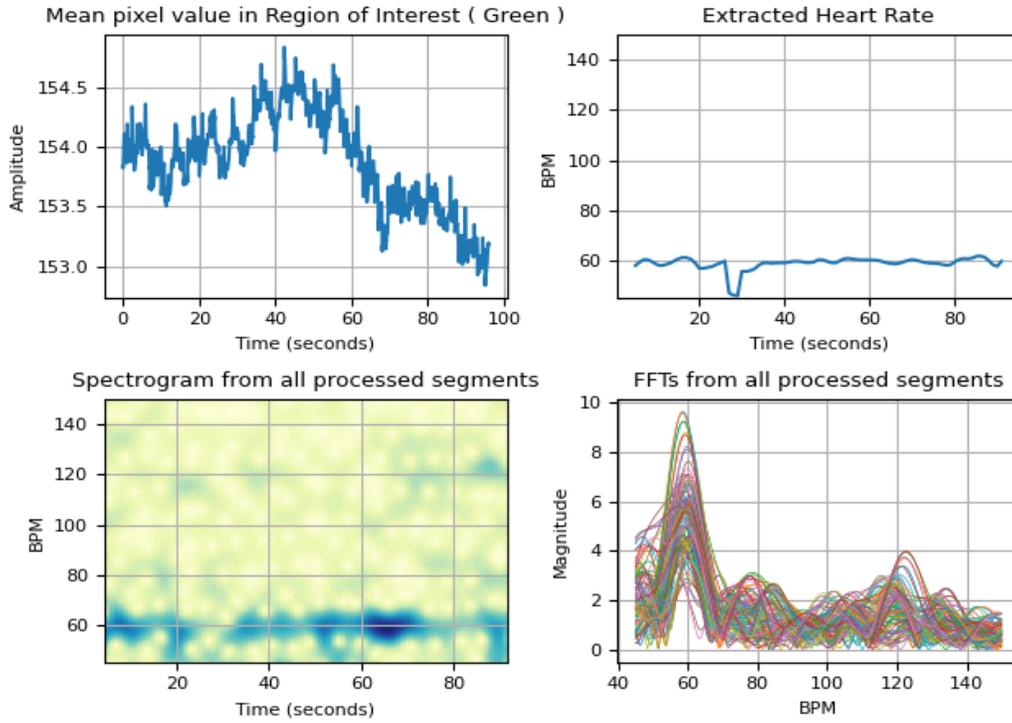


Figure 15. VM algorithm results for condition C4 (sitting, with mask), green, 256x256 ROI from [10]

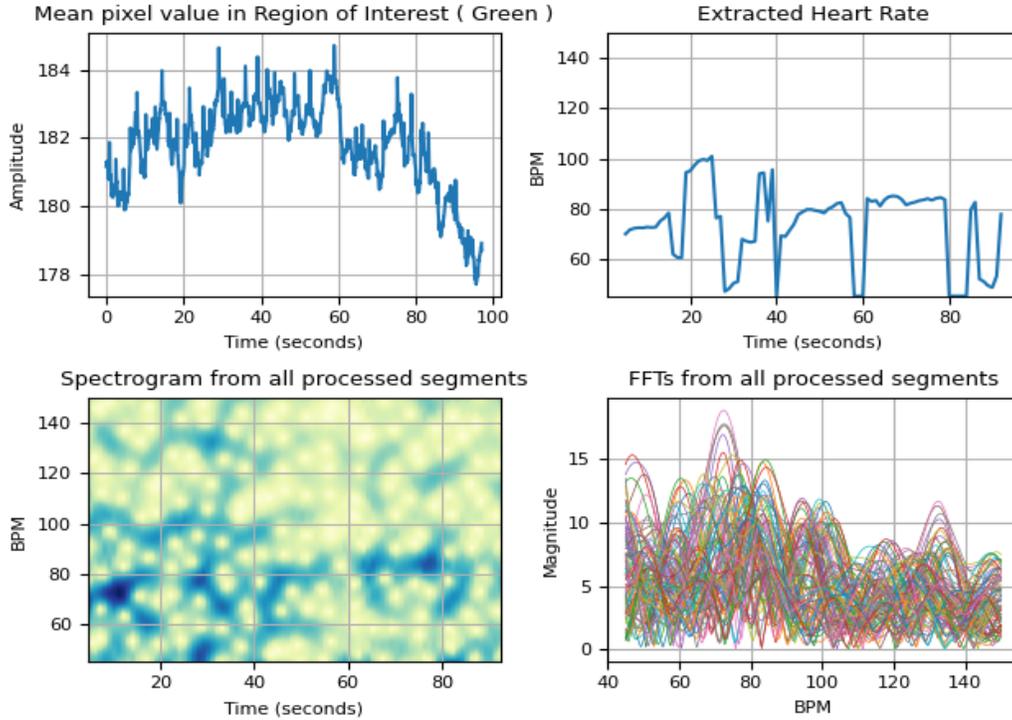


Figure 16. VM algorithm results for condition C5 (standing, no mask), green, 256x256 ROI from [10]

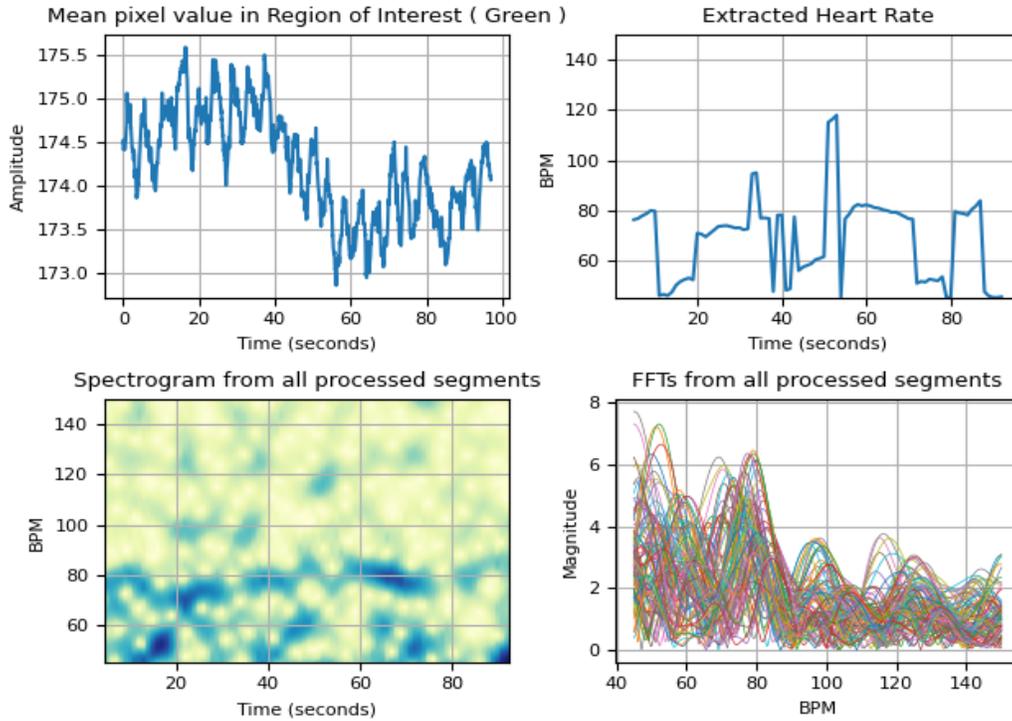


Figure 17. VM algorithm results for condition C8 (standing, with mask), green, 256x256 ROI from [10]

Next, the effect of another input parameter of the VM algorithm is explored, which is the window segment length.

4.2 Effect of Window Segment Length on the Video Magnification Algorithm.

Some of this section's contents are based on the paper in review at the 2022 IEEE Sensors and Applications Symposium titled *Effects of Lighting and Window Length on Heart Rate Assessment through Video Magnification* [11].

Previous works researching the Video Magnification techniques for heart rate detection have used different window segment lengths in their analysis to measure heart rate. These segment lengths included 4 seconds [58], 10 seconds [47], and 30 seconds [48]. However, no work has previously compared the effect of different window segment lengths on the accuracy of the VM algorithm in detecting heart rate. This section reports on the results of varying the window segment lengths from 10 seconds to 20 seconds to 30 seconds with 90% overlap on 19 subjects from the NRC video

database described in the Method and Experimental Setup chapter.

4.2.1 Analysis Details

For this analysis, the NRC testing condition of sitting with mask off (condition C1) was used for all test subjects to reduce the effect of motion on the VM algorithm such as observed in the standing conditions. The same RGB Panasonic camera and the same LED lighting was used as the ones in the analysis results in the previous chapter. 19 out of the 22 participants in the NRC video collection were analyzed, resulting in a subject pool of varying skin tones, age and gender. The region of interest in the analysis was chosen to be of size 64x64 centered on the forehead.

Three participants out of the 22 were excluded from the analysis for the following reasons: one subject had clinically abnormal heart rhythm, another participant had hair occluding their forehead and was removed as they could not have a ROI consistent with the rest of the participants, and the third participant was removed due to missing data during the video collection.

A 90 second video corresponding to each one of the 19 subjects was analyzed three times using the Video Magnification algorithm. In each processing time, all input parameters (ROI location, video length, bandpass filter parameters: 50-150 BPM, and FFT size) were kept constant except for the window segment length, which was varied between 10, 20 and 30 seconds respectively. The measured heart rate obtained from the VM algorithm was then compared with each subject's ground truth heart rate obtained from the pulse oximeter using the percent mean absolute error metric.

Lastly, the average, standard deviation and maximum of the percent mean absolute error for the three window segment lengths were computed and compared. The green colour channel was the focus of the analysis due to it having the best signal to noise ratio as mentioned previously.

4.2.2 Effect of Three Window Segment Length: 10s, 20s, 30s

Table 5 contains the percent mean absolute errors for each of the 19 subjects for the three window segment lengths: 10, 20 and 30 seconds. The subjects are labeled S1-S19. Table 6 contains the average, standard deviation, and maximum of each column in Table 5.

Subject S4 moved during the video capture process and the video contained an episode of sneezing at second 38 of the recoding. The results of the VM algorithm for this subject using a window segment length of 10s can be seen in Figure 18. Due to this motion, the percent mean absolute error is high for all three window segment lengths as shown in Table 5 on the next page.

Subject S18 had the darkest skin tone out of the 19 subjects and he also had the highest percent mean absolute error for all three window segment lengths as shown in Table 5. This decrease in the VM algorithm performance with dark skin tones is not surprising as the literature had confirmed this limitation previously [28]. The VM algorithm results for subject S18 using a 10s window can be seen in Figure 19.

Subject S11 also had a high percent mean absolute error as shown in Table 5 for all three window segment lengths. The subject moves their head and blinks repeatedly in the video (although the ROI is centered on the forehead only and does not contain the eyes). The VM algorithm results for subject S11 can be seen in Figure 20, Figure 21 and Figure 22 for the three window segment lengths 10s, 20s and 30s respectively. The true heart rate of 80-90BPM is seen in the spectrogram plot of Figure 21 and Figure 22 while it was not previously visible in Figure 20.

Motion is still decreasing the accuracy of the VM algorithm and resulting in a high percent mean absolute error, but that error was significantly reduced with increasing window segment length from 20.48% in 10s windows down to 9.52% in 30s windows.

Table 5. Summary of the percent mean absolute error of the 19 subjects for three window segment lengths: 10s, 20s and 30s.

Subject Label:	Percent Mean Absolute Percent Error		
	10s	20s	30s
S1	3.10	2.20	1.76
S2	1.51	1.13	1.18
S3	0.68	0.70	0.92
S4	11.61	15.60	19.11
S5	3.38	1.15	0.99
S6	4.19	6.70	6.02
S7	1.56	1.89	2.27
S8	4.55	1.22	2.03
S9	2.16	1.03	1.39
S10	1.15	1.96	2.15
S11	20.48	15.58	9.52
S12	8.03	1.25	1.02
S13	5.66	5.37	1.54
S14	0.54	1.54	2.26
S15	0.57	1.46	2.26
S16	3.89	1.83	2.42
S17	1.52	2.15	2.44
S18	20.71	19.46	19.19
S19	13.58	12.70	9.87

Table 6. Average, standard deviation and mean of the percent mean absolute error of the 19 subjects for the three window segment lengths: 10s, 20s and 30s.

	10s	20s	30s
Average	5.73	4.99	4.65
Standard Deviation	6.19	5.88	5.59
Maximum	20.71	19.46	19.19

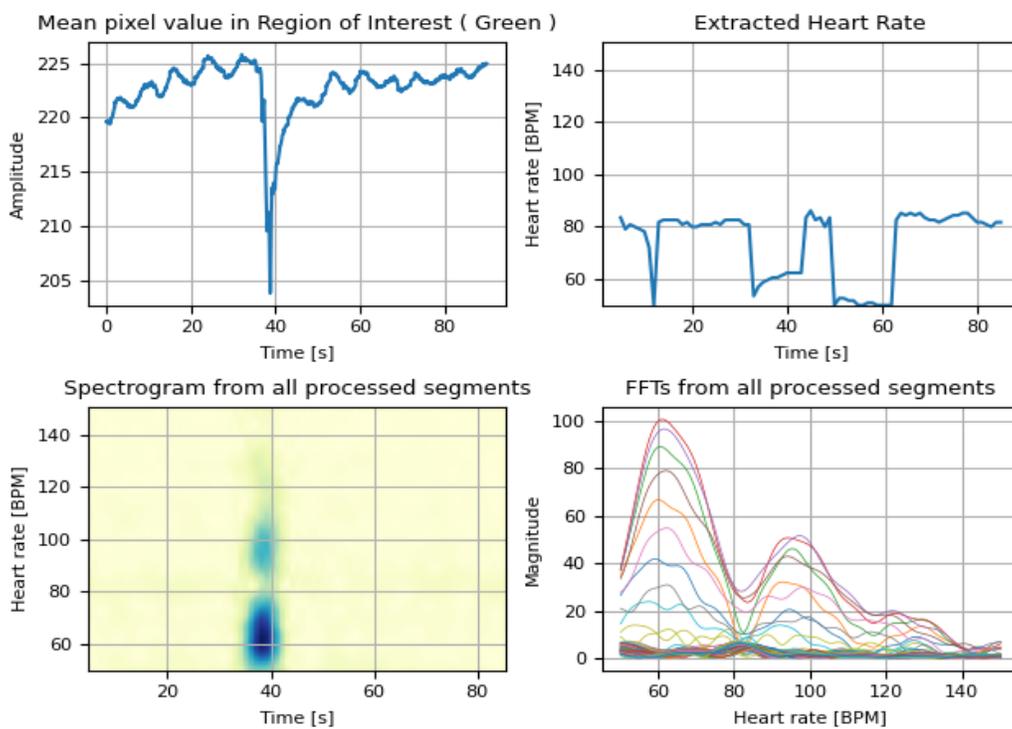


Figure 18. The VM algorithm results for subject S4 using a 10s window segment length

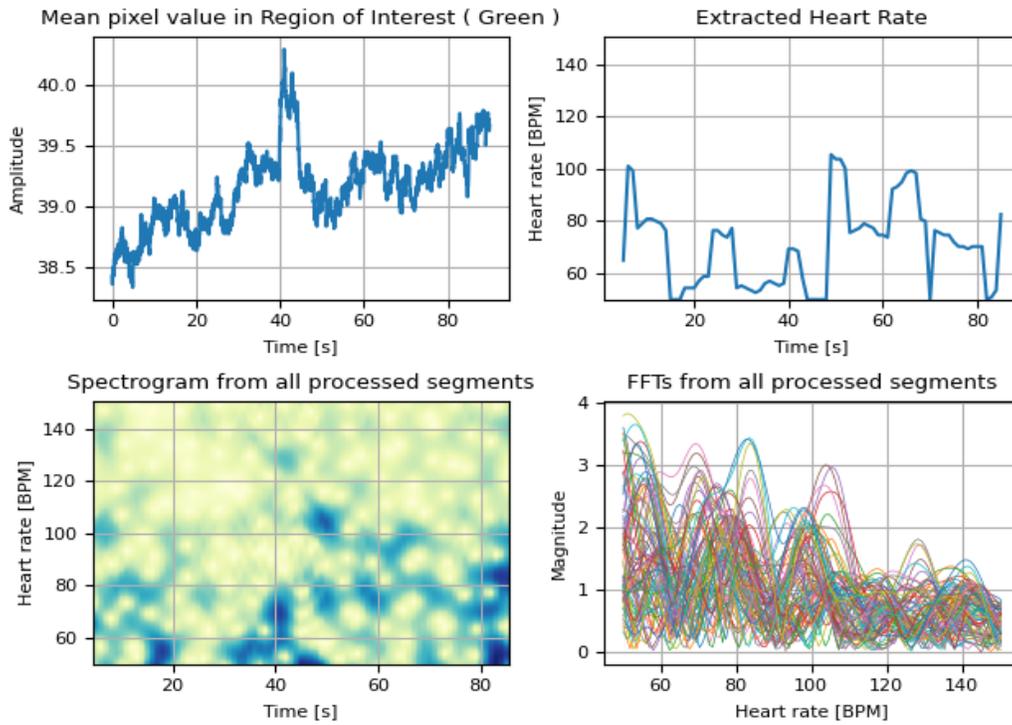


Figure 19. The VM algorithm results for subject S18 using a 10s window segment length

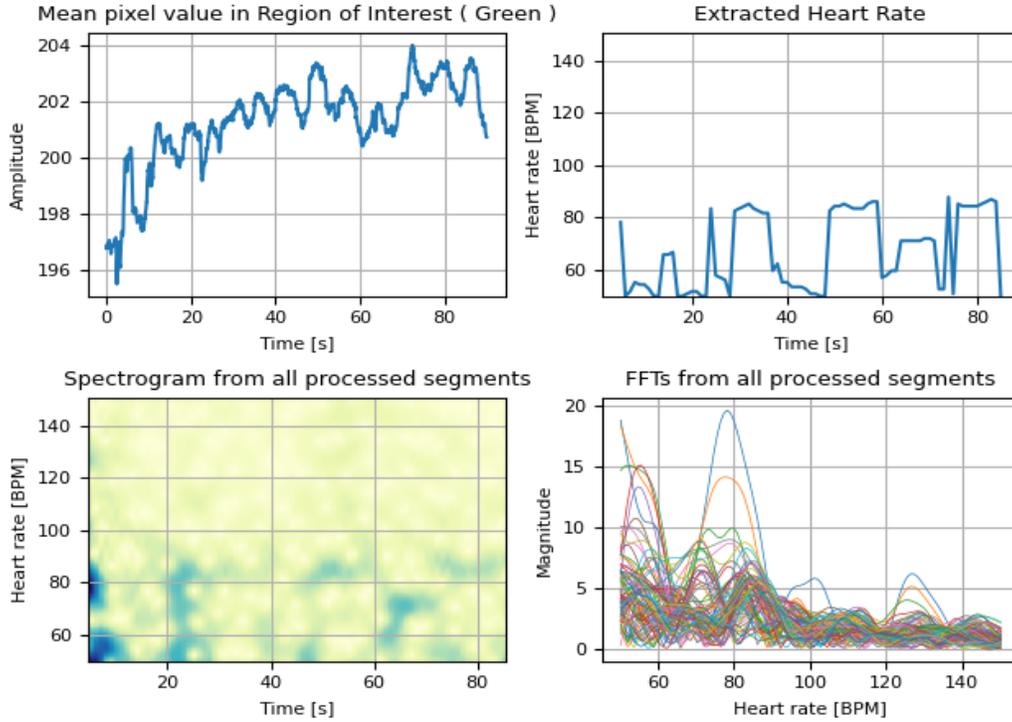


Figure 20. The VM algorithm results for subject S11 using a 10s window segment length

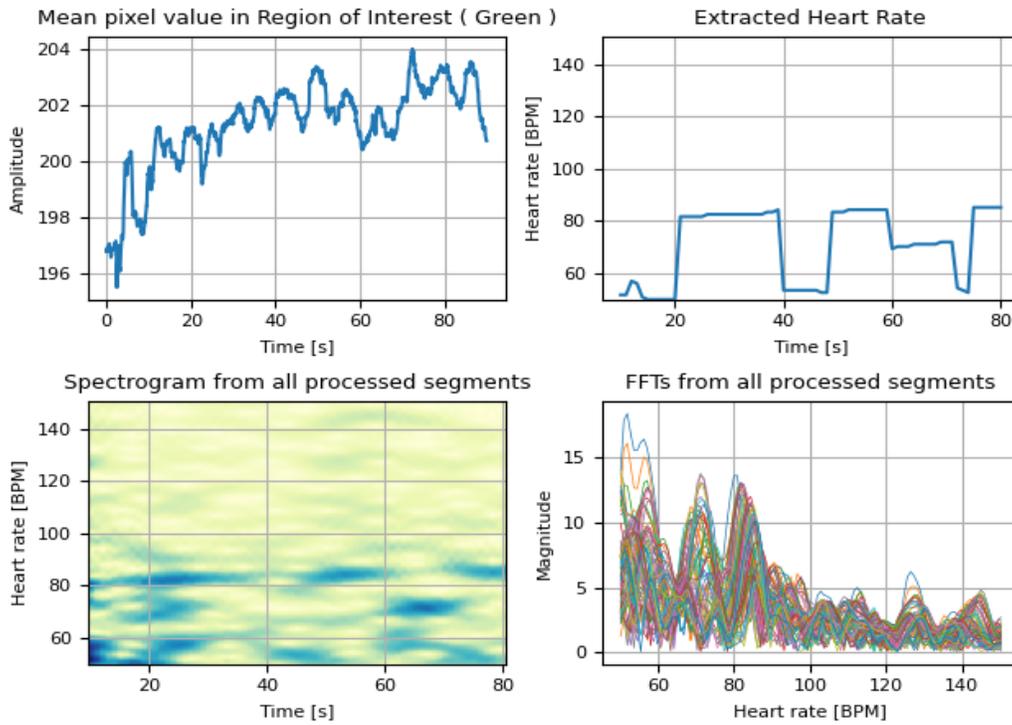


Figure 21. The VM algorithm results for subject S11 using a 20s window segment length

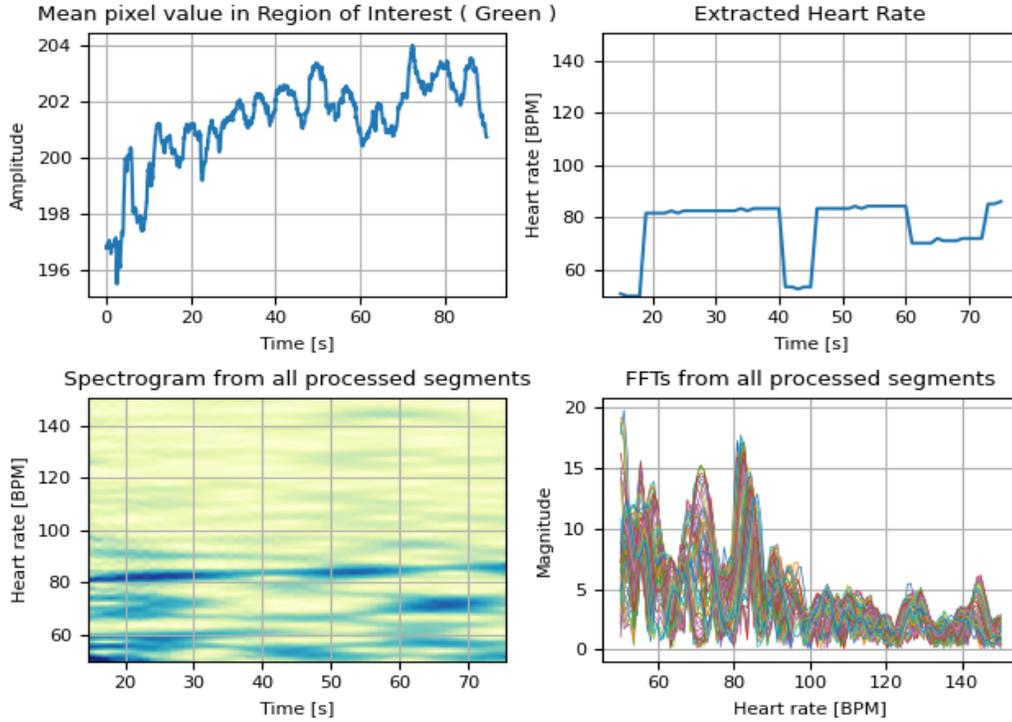


Figure 22. The VM algorithm results for subject S11 using a 30s window segment length

Similar to subject S11, subject S19 had a fair skin tone, and a relatively high percent mean absolute error that decreased with increasing window segment length as shown in Table 5. The subject also displayed head motion, although it was less frequent than subject S11. Figure 23, Figure 24 and Figure 25 show the VM algorithm results for subject S19 for the three window segment lengths 10s, 20s and 30s respectively. For a true heart rate of 85-90BPM, the spectrogram plots in the three figures contain the true heart rate but the effect from the motion that is seen as a lower frequency component is decreased with increasing window segment length.

The results in Table 5 divide the subjects into two groups. In one group the VM results improve with increasing window segment length and in the other group the results deteriorate. After watching all videos, it was found that the group with improving VM results had visible motion of the head in the videos while the other group was almost completely motionless. When the subject was motionless, the VM algorithm detected with a smaller window segment length (10s) the true heart rate with high accuracy.

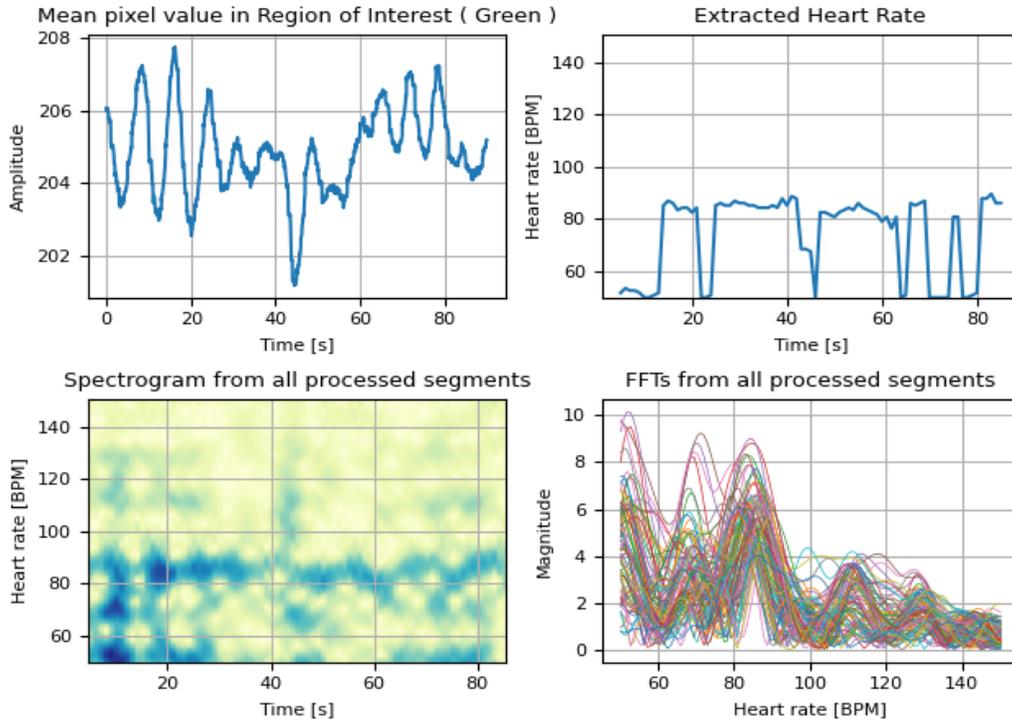


Figure 23. The VM algorithm results for subject S19 using a 10s window segment length

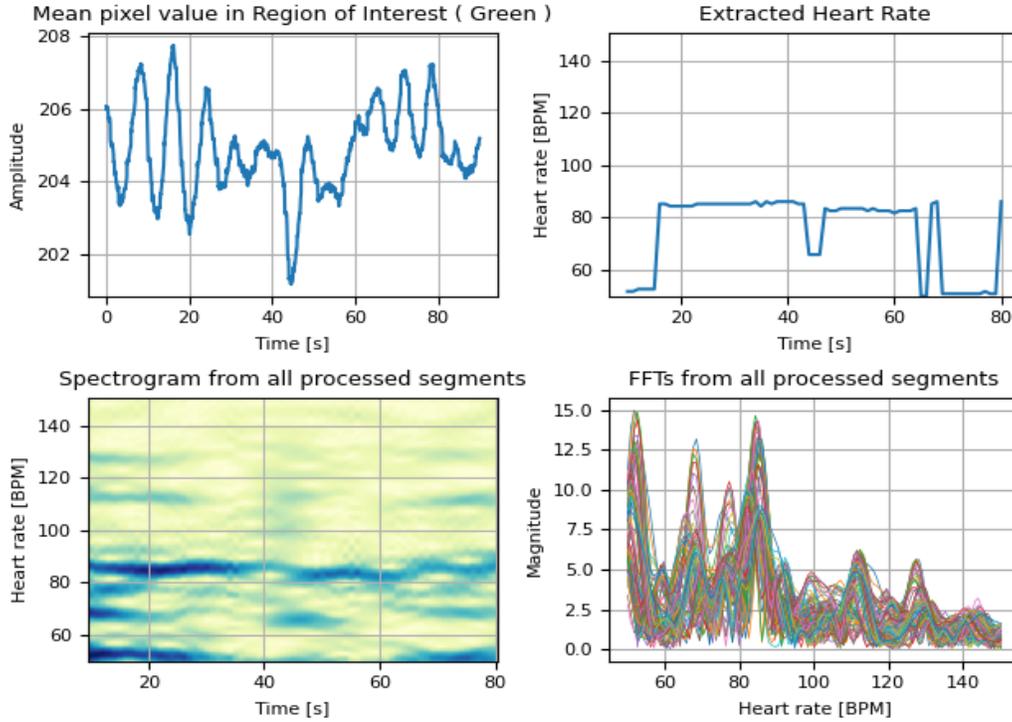


Figure 24. The VM algorithm results for subject S19 using a 20s window segment length

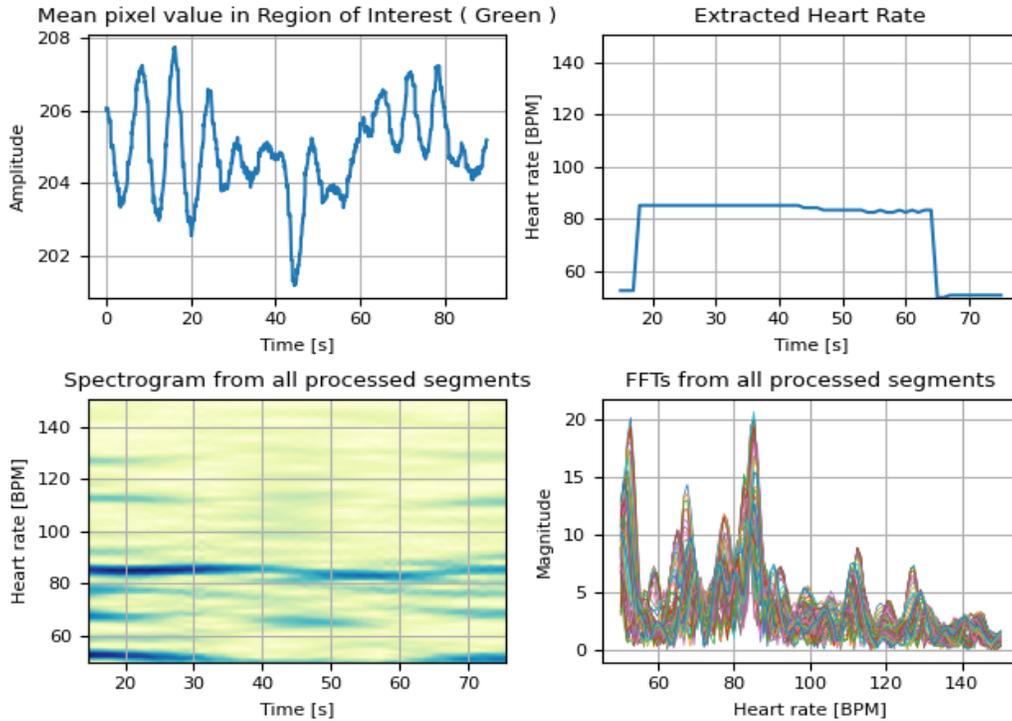


Figure 25. The VM algorithm results for subject S19 using a 30s window segment length

The small changes in the true heart rate were detected and this resulted in very small percent mean absolute error such as the case for S3, S7, S10, S14, S15, and S17 summarized in Table 5. Increasing the window segment length in this condition caused the algorithm to lose sensitivity which resulted in higher percent mean absolute error. This can be seen when comparing the true heart rate reading of subject S14 in Figure 26 to the VM algorithm results obtained using the three window segment lengths 10s, 20s and 30s respectively in Figure 27, Figure 28, and Figure 29. The sensitivity in the heart rate between 45-55 seconds of video time is detected by the VM algorithm with a smaller window segment length (10s) but there is an effect of averaging in the two bigger window segment lengths (20s and 30s).

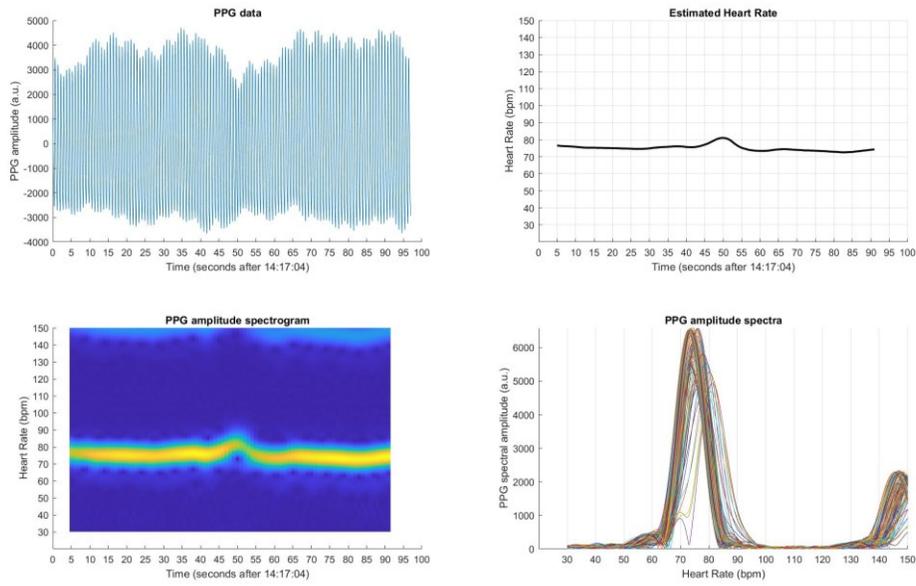


Figure 26. True heart rate from PPG for subject S14

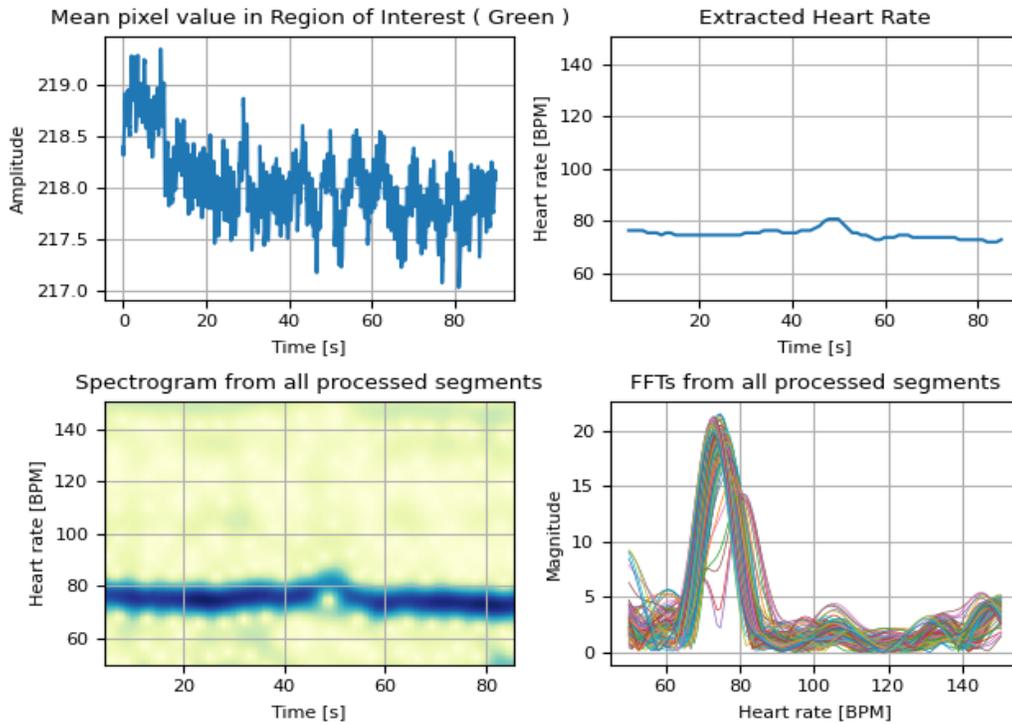


Figure 27. The VM algorithm results for subject S14 using a 10s window segment length

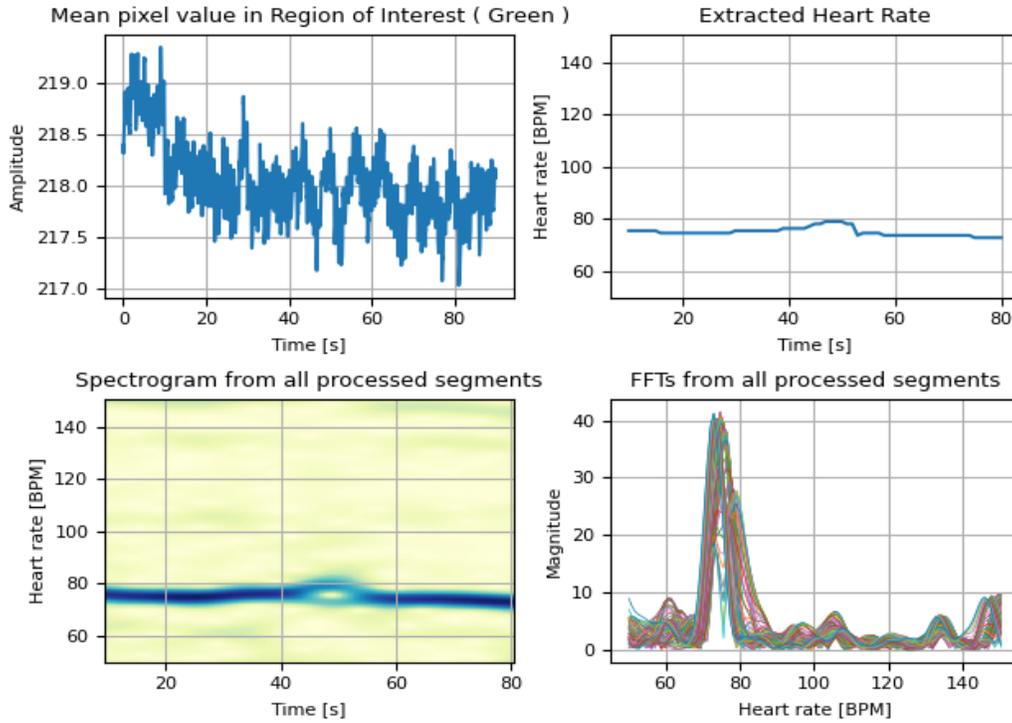


Figure 28. The VM algorithm results for subject S14 using a 20s window segment length

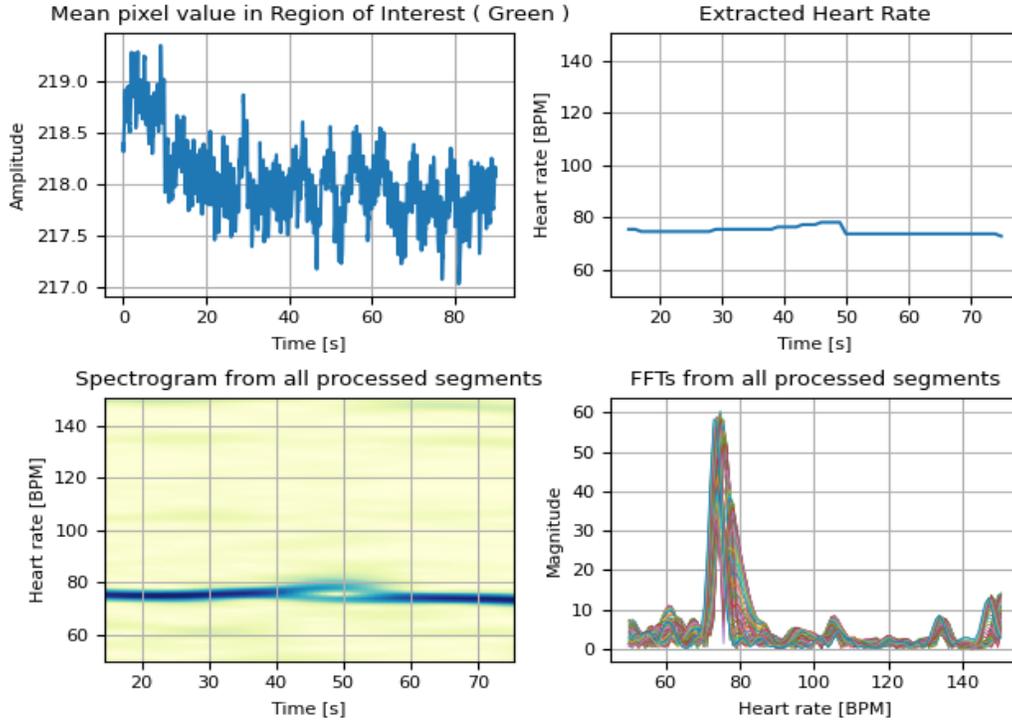


Figure 29. The VM algorithm results for subject S14 using a 30s window segment length

4.3 Conclusion

In conclusion, heart rate can be detected by the VM algorithm from very small ROIs but the accuracy of detection improves with increasing ROI size.

Additionally, in the presence of motion, increasing the window segment length used to process the video results in an improvement in the HR detection accuracy via VM. On the other hand, with a lack of motion, a window segment length of 10 seconds can detect heart rate with high accuracy using the VM algorithm.

Chapter 5: Effect of Noise

In this chapter, the effect of noise from three different sources on the Video Magnification algorithm performance is detailed. The process of taking a video and measuring heart rate from it through Video Magnification techniques involves many parameters including the camera, the lighting condition, the skin tone of the subject, and the algorithm and coding as illustrated in Figure 30. Noise may arise from each of these stages, and as a result it affects the performance of the VM algorithm. The three investigated noise sources in this chapter are quantization noise, noise from modern camera systems equipped with software enhancements, and noise from illumination level from different lighting sources respectively.

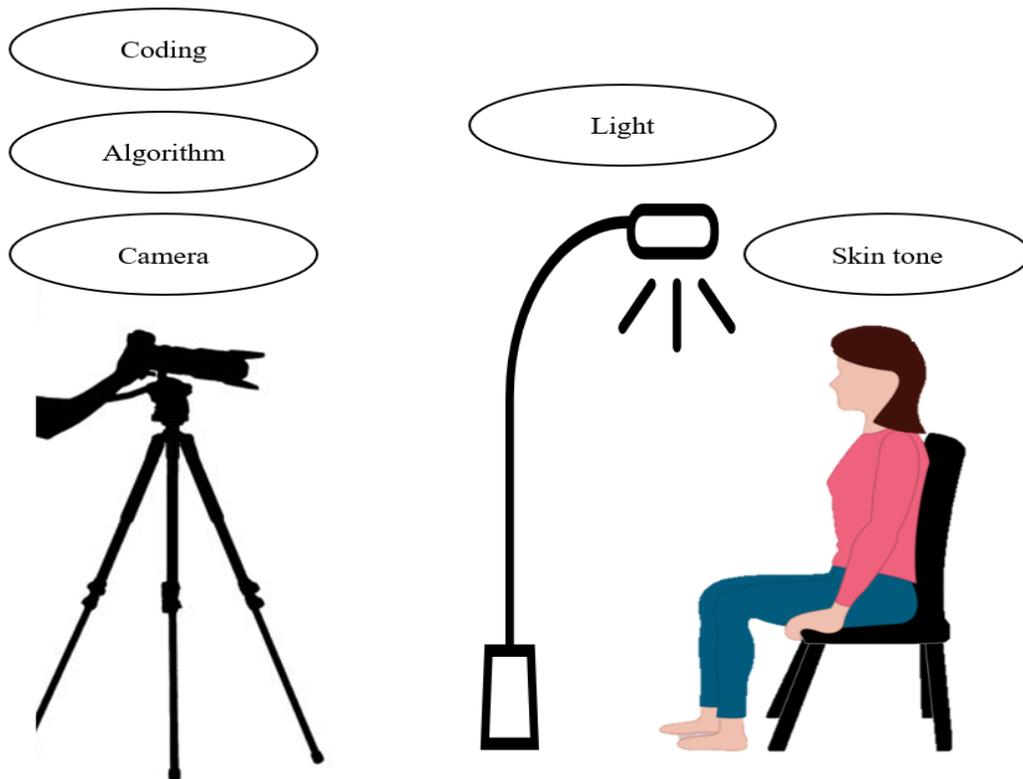


Figure 30. Sources of noise in remote heart rate measurement from video

5.1 Effect of Quantization Noise on the Video Magnification Algorithm

This section's contents are based on the paper accepted for publication at the 2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) titled *The Effect of Noise on Contactless Heart Rate Measurement using Video Magnification* [12].

The Video Magnification algorithm performance was investigated using artificially made videos with large and small signal amplitudes at frequencies resembling heart rate. Additionally, with small signal amplitudes, the VM algorithm performance was investigated in the presence and absence of white Gaussian noise.

Previous work had investigated the Video Magnification performance with noise using artificially generated thermal videos containing known Gaussian noise levels [37]. Another work had investigated the noise added to captured RGB video on the VM algorithm, however the noise in the source videos was unknown [33]. Nevertheless, no work had previously investigated the effect of small signal amplitude in the presence and absence of noise on the VM algorithm performance using colour videos with quantified noise. This chapter reports on the investigation of the smallest signal amplitude at which the VM algorithm correctly detects the heart rate frequency from in the presence and absence of noise. This investigation gives insights into the quantization noise of the system.

The artificially generated videos described in Chapter 3 were used to generate the results of this section. As previously mentioned, the standard representation of values in videos is 8-bit integers, therefore when the signal is converted to integers, the value of k is inter-related with the quantization levels. Since the signal range is 0 to 1 prior to the conversion, then the quantization spacing is $1/(2^8)$, which equals 0.00391. Therefore, levels between ± 0.00196 ($0.00391/2$) within a quantization level would be rounded to that level.

5.1.1 Artificial Video Processing

In total, there were 5 signal amplitudes for the large signal amplitude group, and 4 signal amplitudes for the small signal amplitude group. Additionally, there were 6 noise levels and 4 heart rate frequencies simulated. This resulted in 20 videos that were processed for the large signal amplitudes, 16 videos that were processed for the small signal amplitudes without noise, and 96 videos that were processed for the small signal amplitudes with 6 noise levels and 4 heart rate frequencies. All videos were then run through the Video Magnification algorithm. The input parameters were set to 10s video length, 10s window segment length, bandpass parameters of 0 to 200BPM (which is different from what was previously used so far, due to the nature of this experiment) and FFT size of 2048.

To measure the performance of the VM algorithm in the presence of noise with the small signal amplitudes, a figure of merit comparing the magnitude of the computed FFT at the known correct heart rate in the video to the magnitude of the highest incorrect FFT peak in the graph was used according to Equation 5.

Equation 5. The ratio used for the figure of merit to quantify the VM algorithm performance

$$\text{FFT Ratio} = \frac{\text{Magnitude of correct FFT peak}}{\text{Magnitude of largest incorrect FFT peak}}$$

An FFT Ratio less than 1 would indicate that the VM algorithm would make an error with the frequency detection.

5.1.2 Effect of Large and Small Signal Amplitude in the Absence of Noise

For the large signal amplitudes, taking the generated artificial videos at 53BPM frequency, the results of the VM algorithm for $k = 0.05-0.01$ signal amplitudes can be seen in Figure 31-Figure 35 respectively. The signal appears sinusoidal in all five figures and the computed FFT has a peak at the true heart rate. No other frequencies (noise) are detected from the artificial videos by the VM

algorithm, which confirms the correct coding and implementation of the algorithm. Figure 35 that corresponds to the smallest signal amplitude in the large signal range shows the beginning of a distortion in the sinusoidal signal and the heart rate's third harmonic (159BPM) starts to appear. Figure 36 compares the average magnitude of the peak FFT at the correct heart rate to the large signal amplitude k for the four different heart rates. The graph shows a linear relationship between the k value and the magnitude of the peak FFT, indicating that the magnitude of the FFT increases linearly with increasing k between 0.01 and 0.05.

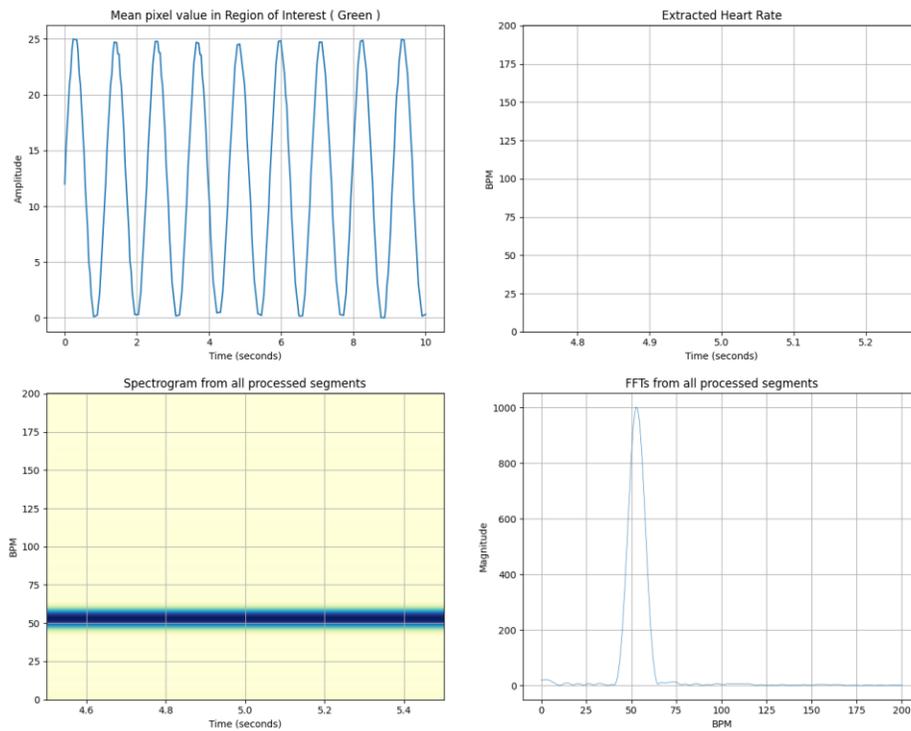


Figure 31. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.05

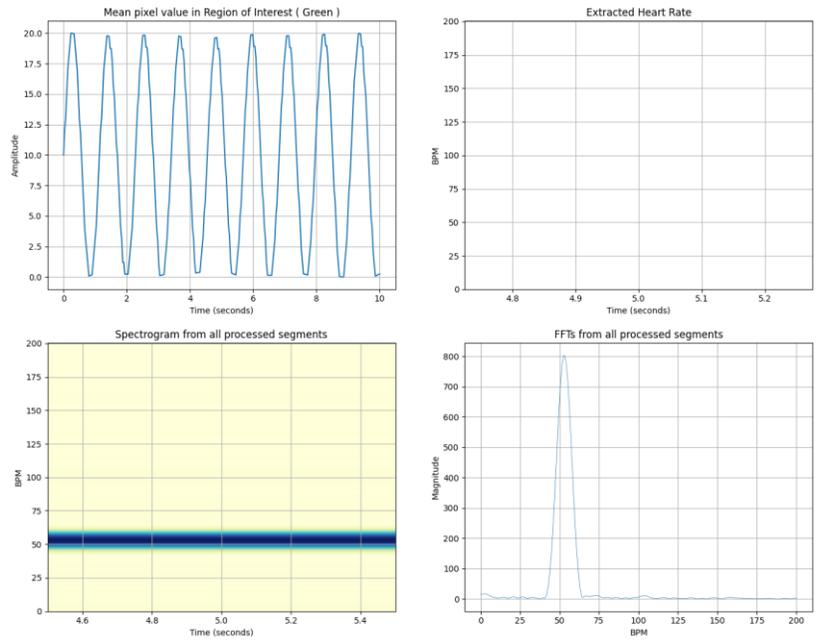


Figure 32. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.04

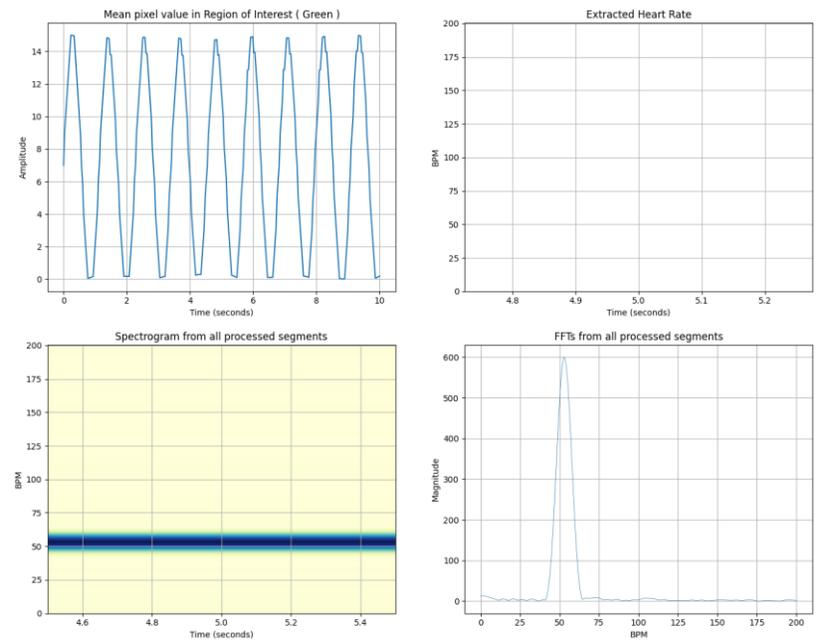


Figure 33. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.03

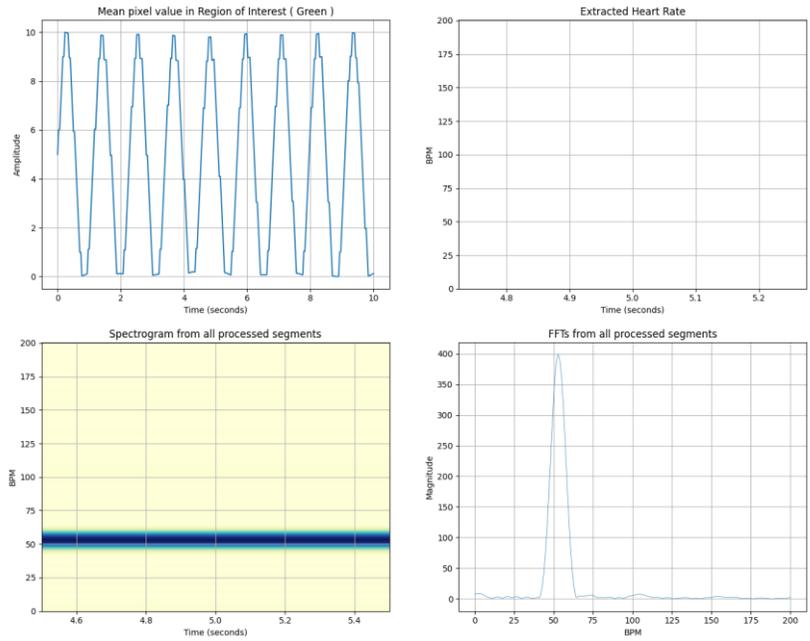


Figure 34. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.02

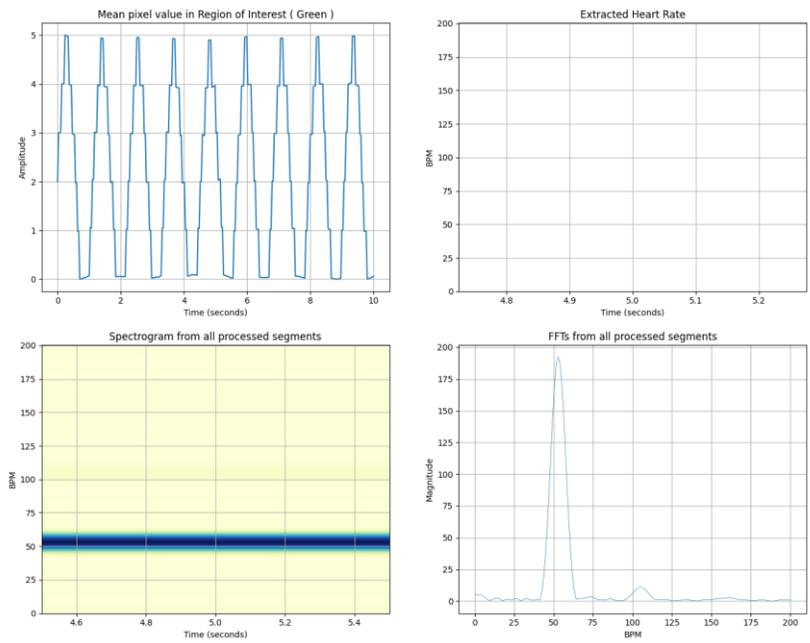


Figure 35. VM algorithm results for the 53BPM artificial video at large signal amplitude 0.01

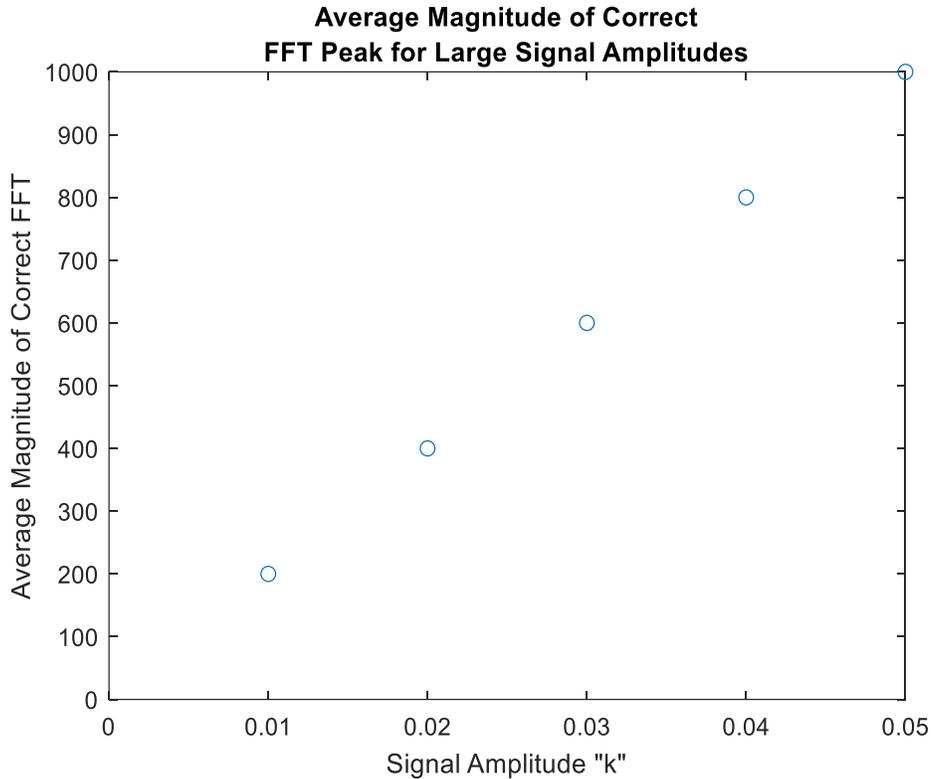


Figure 36. Average of the peak FFT magnitude at the true heart rate for $k = 0.01-0.05$ from [12]

For the small signal amplitudes without noise, taking the 53BPM generated artificial videos, the VM algorithm results for $k = 0.005-0.002$ signal amplitudes can be seen in Figure 37-Figure 40 respectively. Figure 41 shows the VM algorithm results for $k = 0.00196$. In Figure 37-Figure 40, the signal's periodicity is still visible in the plots, but the sinusoidal signal shape starts to resemble a square wave. Additionally, the magnitude of the FFT at the heart rate's third harmonic (159BPM) increases with decreasing k value. In Figure 40, the FFT peak at the true heart rate had decreased significantly where it is closely approaching the FFT peak at the third harmonic.

$k = 0.002$ corresponds to an amplitude just above the quantization spacing of an 8-bit video and it is the smallest signal amplitude at which the VM algorithm was able to correctly detect the true heart rate in the signal in the absence of noise. The presence of the third harmonic is an expected behavior since the sine wave's mean pixel value is increasingly more square-wave like until finally

being reduced to an alternating impulse train with an amplitude of ± 1 pixel. This indicates that the system is approaching the limits of quantization. Additionally, the Fourier series of this impulse train contains the odd harmonics of the sine wave frequency. At $k = 0.00196$, the VM algorithm does not detect a heart rate in the signal as the signal becomes a constant as shown in Figure 41. This is in line with the expected behavior of the system for 8-bit integer quantization levels. Figure 42 compares the average magnitude of the peak FFT at the correct heart rate to the small signal amplitude k including $k = 0.00196$ for the four different heart rates. The figure shows an exponential relationship between the k value and the magnitude of the peak FFT, indicating that the magnitude of the FFT increases while approaching an exponential limit with increasing k between 0.002 and 0.005.

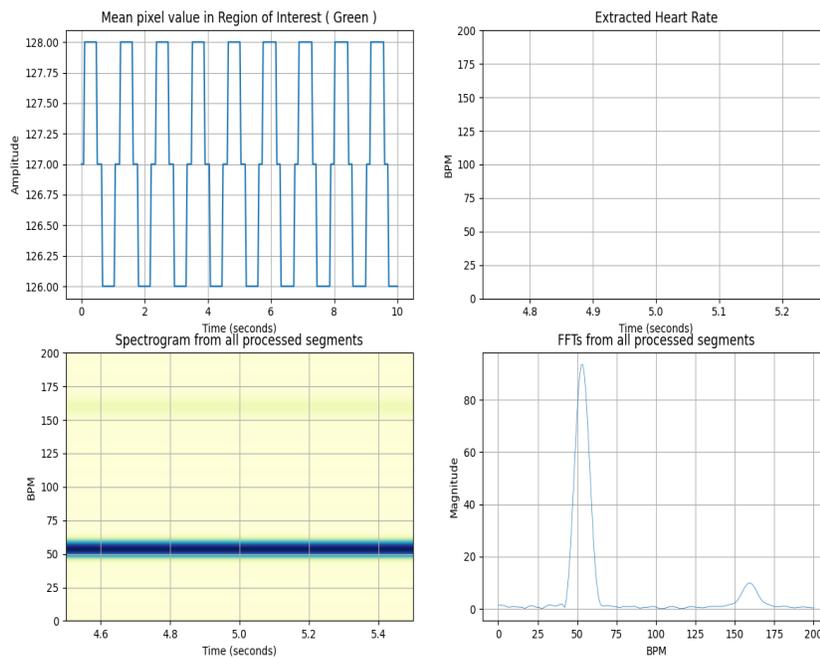


Figure 37. VM algorithm results for the 53BPM artificial video at small signal amplitude 0.005

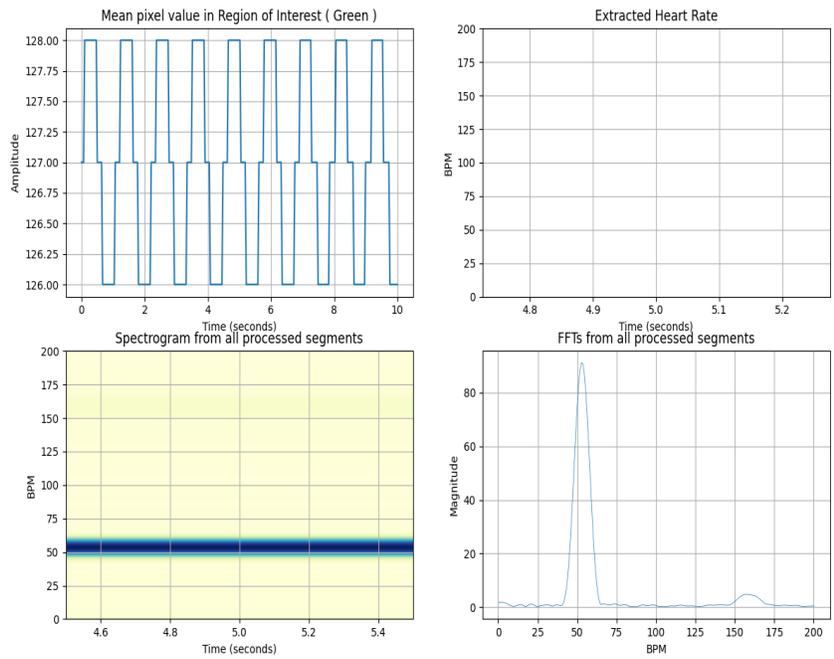


Figure 38. VM algorithm results for the 53BPM artificial video at small signal amplitude 0.004

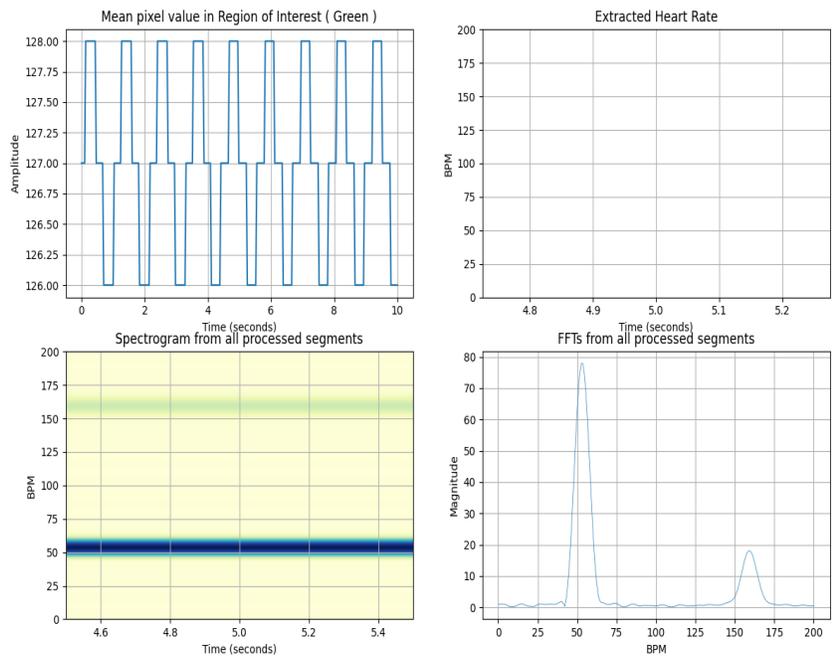


Figure 39. VM algorithm results for the 53BPM artificial video at small signal amplitude 0.003 from [12]

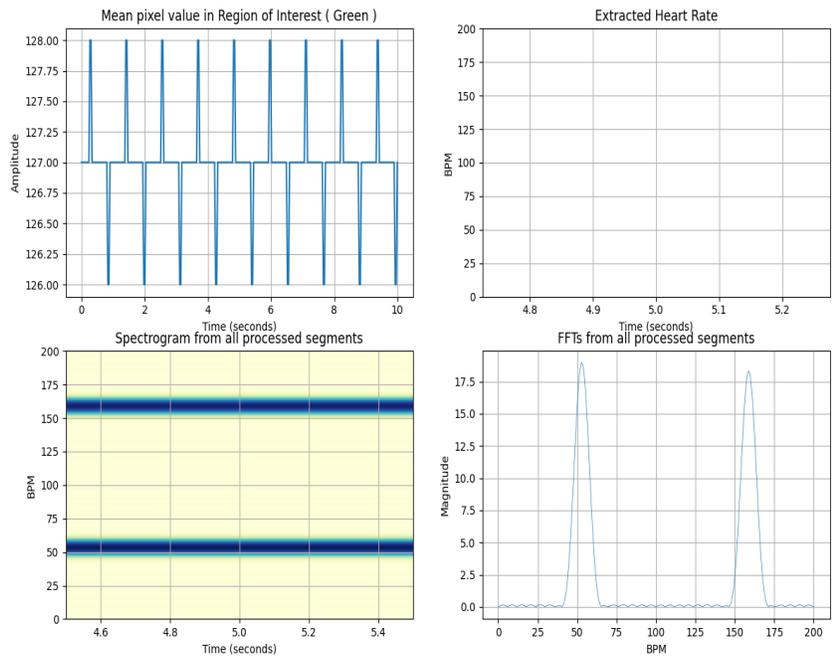


Figure 40. VM algorithm results for the 53BPM artificial video at small signal amplitude 0.002 from [12]

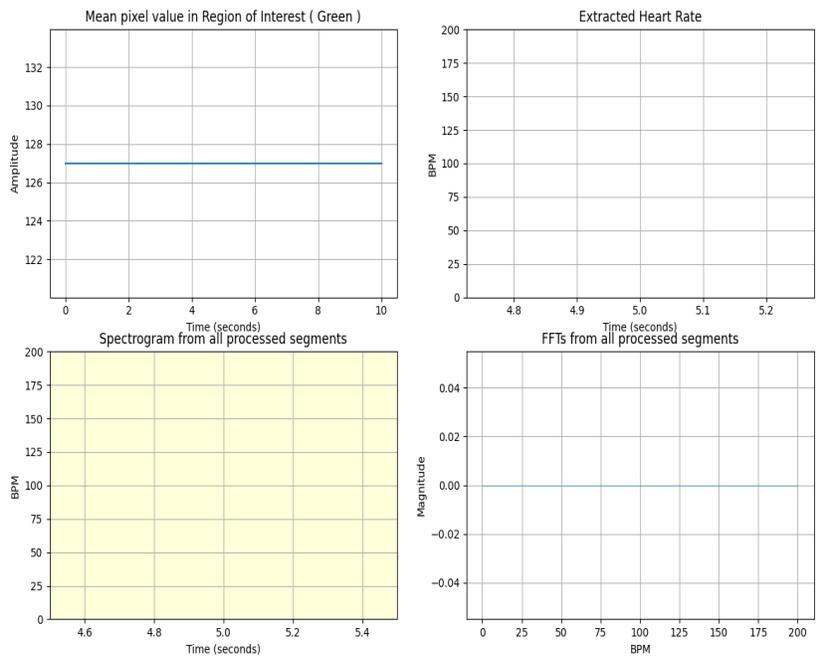


Figure 41. VM algorithm results for the 53BPM artificial video at signal amplitude 0.00196

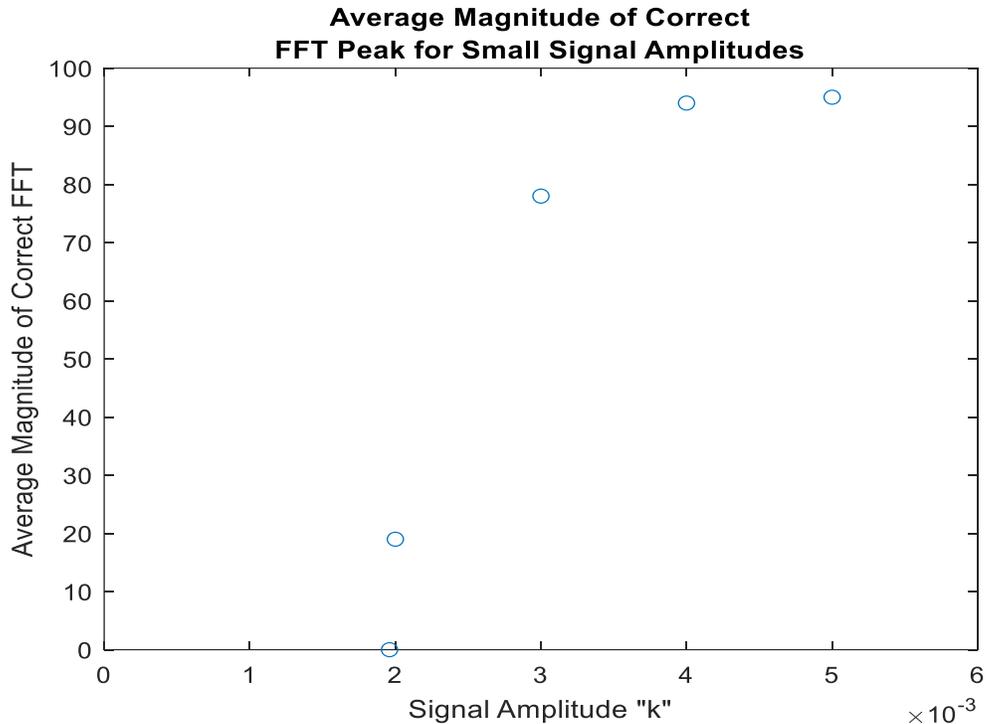


Figure 42. Average of the peak FFT magnitude at the true heart rate for $k = 0.002-0.005$ from [38]

5.1.3 Effect of Small Signal Amplitude in the Presence of Noise

The 6 noise levels that were added could be further divided into two groups, small noise ($m = 5, 10, 15$) and large noise ($m = 20, 25, 30$). Figure 43 shows the average and standard deviation of the FFT Ratio metric vs. the noise level m for the different small signal amplitudes.

A threshold of 2 for the FFT Ratio was setup to compare the VM algorithm performance in the presence of noise. If the FFT Ratio is above 2, then the VM algorithm can detect the true heart rate in the signal accurately with confidence, and if it is less than 2, then the performance has degraded.

At the small noise levels, the FFT Ratio for all small signal amplitudes had an average above 2. This indicates that the VM algorithm could accurately detect the true heart rate in the signal in the presence of small amount of noise ($m=5, 15, 20$).

As the noise increased to the large noise levels, at $m = 20$, all signal amplitudes except $k = 0.002$

had a FFT Ratio average above 2. At $m=25$, only the highest $k=0.005$ had a FFT Ratio average above 2. Finally, at $m=3$, none of the signal amplitudes had a FFT Ratio average above 2. This information indicates that in the presence of large amounts of noise ($m=30$), the VM algorithm's performance degrades at detecting the true heart rate in the signal.

Additionally, in the presence of large amounts of noise, the VM algorithm's performance is worse for smaller signal amplitudes. Nonetheless, for the $k=0.002$ case, the addition of a small amount of noise ($m=15$) has improved the VM algorithm performance at detecting the true heart rate as seen in Figure 44 when compared to Figure 40. With the addition of noise at this small signal amplitude, the signal appears to be within ± 2 quantization levels even though it would be expected to be within ± 1 quantization level. This improvement in performance of the VM algorithm could be due to the effect of dithering by the noise increasing the quantization level range.

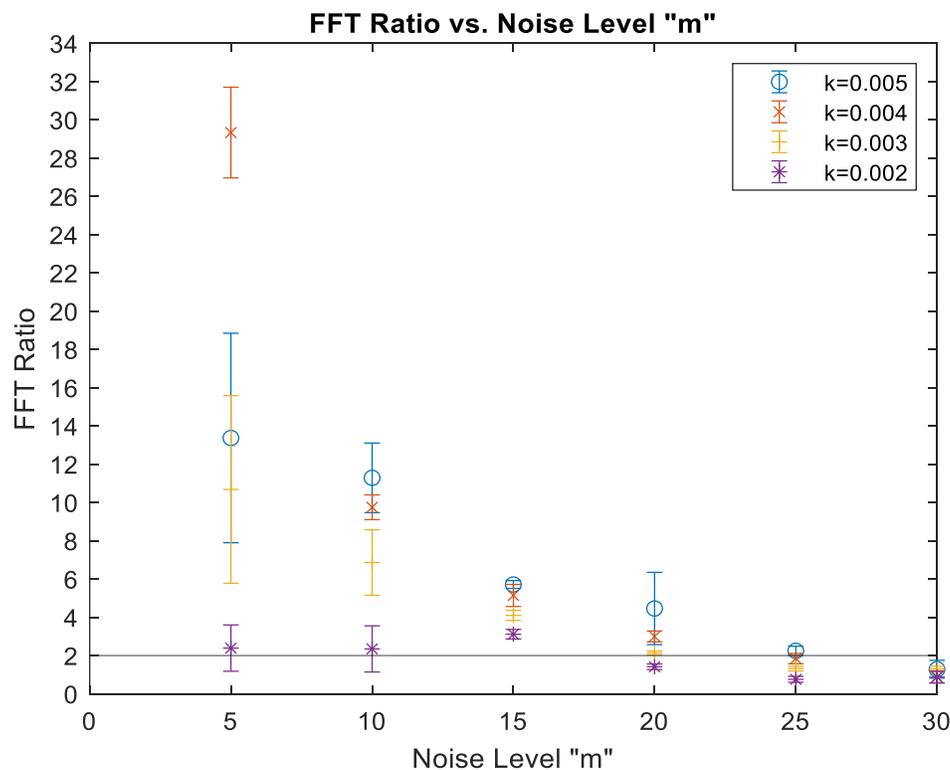


Figure 43. Average and standard deviation of the FFT Ratio for the different noise levels m from [12]

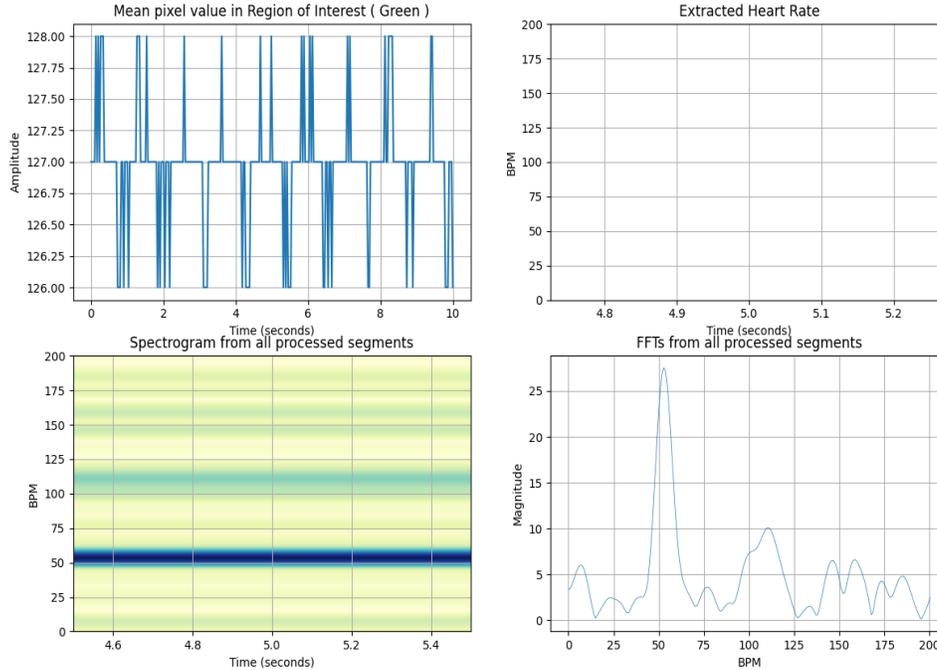


Figure 44. VM algorithm results for the 53BPM artificial video at $k= 0.002$ and $m=15$ from [38]

Next, the effect of modern advanced camera systems equipped with software enhancement features on the performance of the Video Magnification algorithm for heart rate detection is investigated.

5.2 Effect of Camera on the Video Magnification Algorithm.

Research using Video Magnification techniques for measuring heart rate has experimented with lower grade commercial HD webcams and smart phone cameras and had been able to obtain heart rate results with good accuracy. However, with the recent advances in smart phone camera technology, there has been hardware improvements coupled with software and machine intelligence features that apply smart changes to the pixels within video to improve the appearance of the video for general smart phone use. These smart changes may not affect all pixels within a video frame equally. In this section, the smart video features of an iPhone 12 Pro Max are investigated and their effect on the performance of the Video Magnification algorithm for heart

rate detection is reported. The results from the iPhone 12 Pro Max videos using the VM algorithm are compared with the results obtained using a VGA webcam in the same setting. All results are compared against the true heart rate results obtained from a pulse oximeter as described earlier. The videos used to generate the results of this section are the ones detailed in the Self-Capture Video Method.

5.2.1 Video Magnification Performance with VGA Web Camera

The VM algorithm results for the video captured by the VGA camera can be seen in Figure 45 and the true heart rate can be seen in Figure 46. Using a standard VGA camera, the VM algorithm can detect the heart rate of a stationary subject with high accuracy.

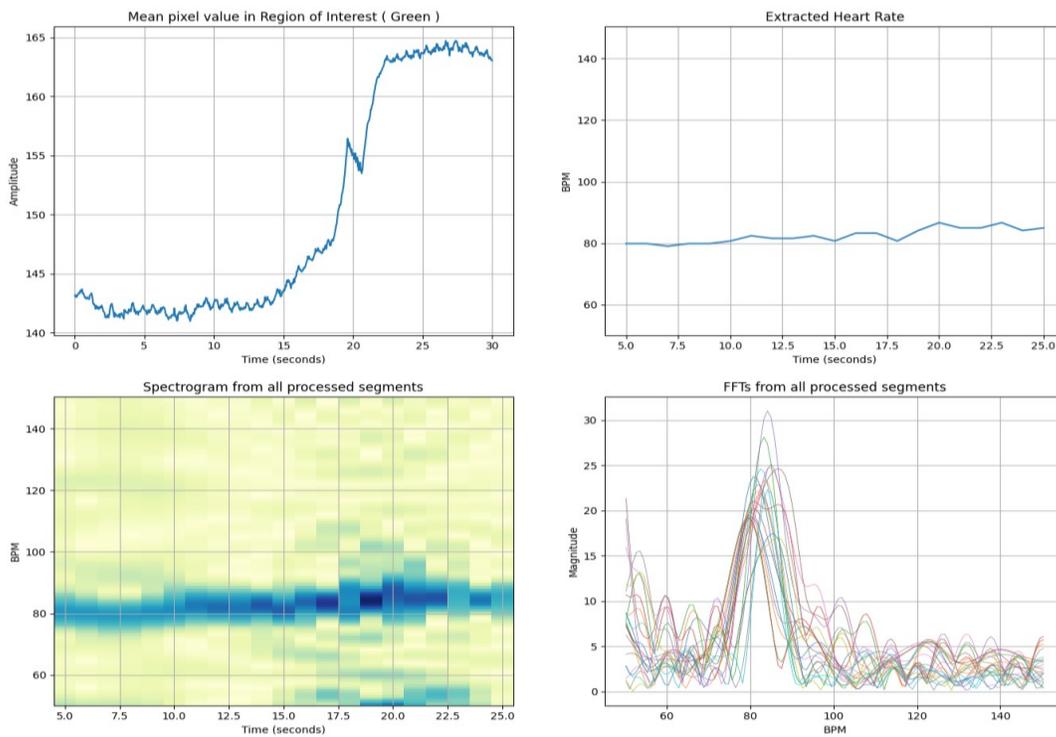


Figure 45. The VM algorithm results for a VGA camera video

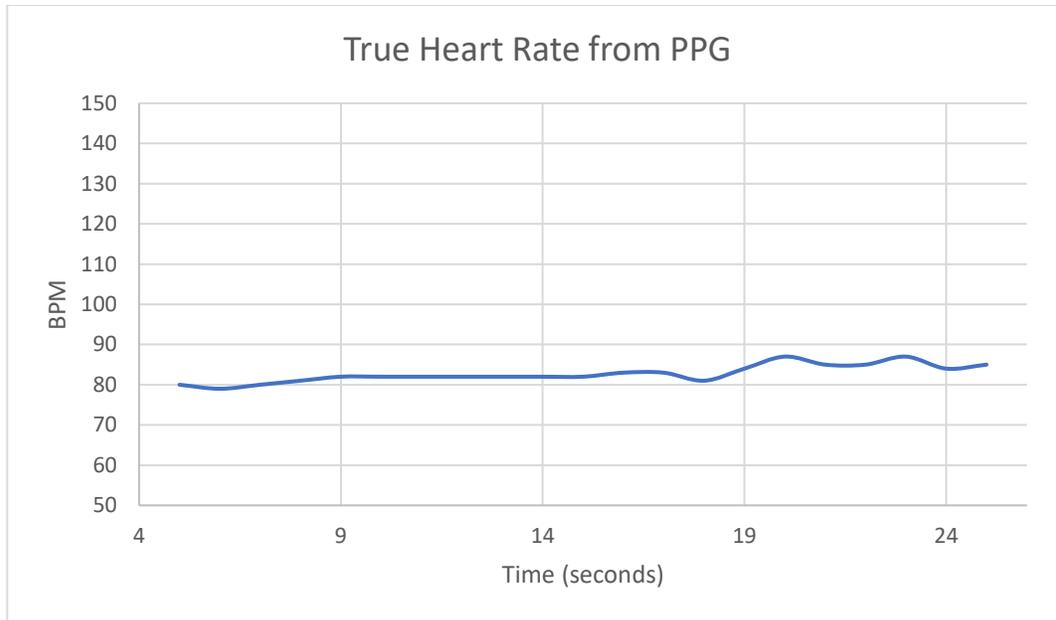


Figure 46. True heart rate results from PPG for VGA video

5.2.2 Video Magnification Performance with iPhone 12 Pro Max Camera

The VM algorithm results for the iPhone 12 Pro Max videos can be seen in Figure 47 through Figure 51 with the true heart rate shown in Figure 52. When using the better-quality lens of the iPhone 12 Pro Max’s wide camera with all features turned off, the VM algorithm could detect the true heart rate from the high-quality camera with high accuracy as shown in Figure 47. Furthermore, The VM algorithm’s performance did not seem to be affected by the HDR smart feature that records in 10-bit high dynamic range compared to the standard 8-bit range as shown in Figure 52. It is important to note however that this is the case with a stationary subject. On the other hand, with the Auto 30FPS smart feature turned on, the VM algorithm results showed a gradual sharp increase in heart rate over the 30 seconds video from 72 to 80BPM as seen in Figure 49 that was not reflective of the true heart rate of the subject. The same trend is seen in Figure 50 and Figure 51 and the gradual increase in measured heart rate appears even sharper for the Auto 30&60FPS and Auto30&60FPS with HDR smart features.

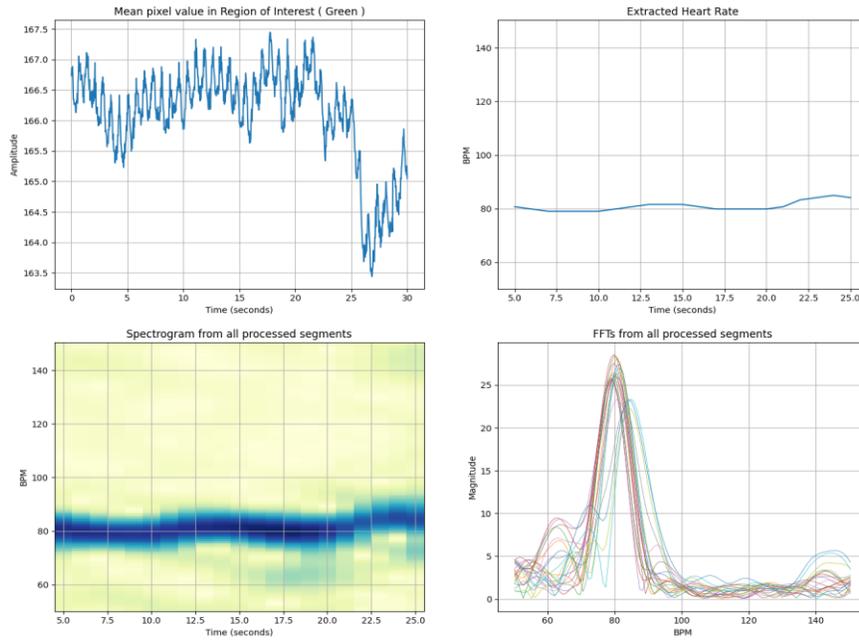


Figure 47. The VM algorithm results for an iPhone 12 Pro Max video with all smart features turned off

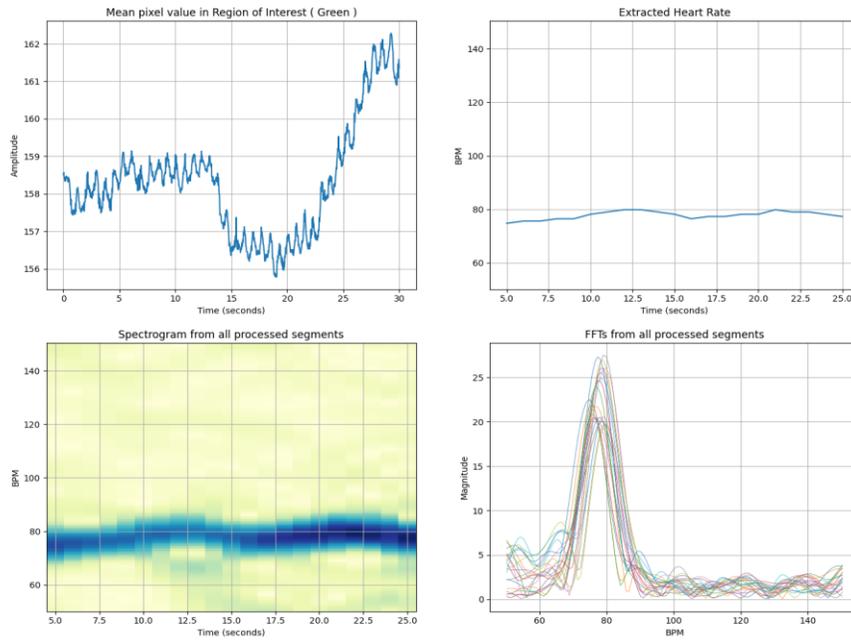


Figure 48. The VM algorithm results for an iPhone 12 Pro Max video with smart HDR feature turned on

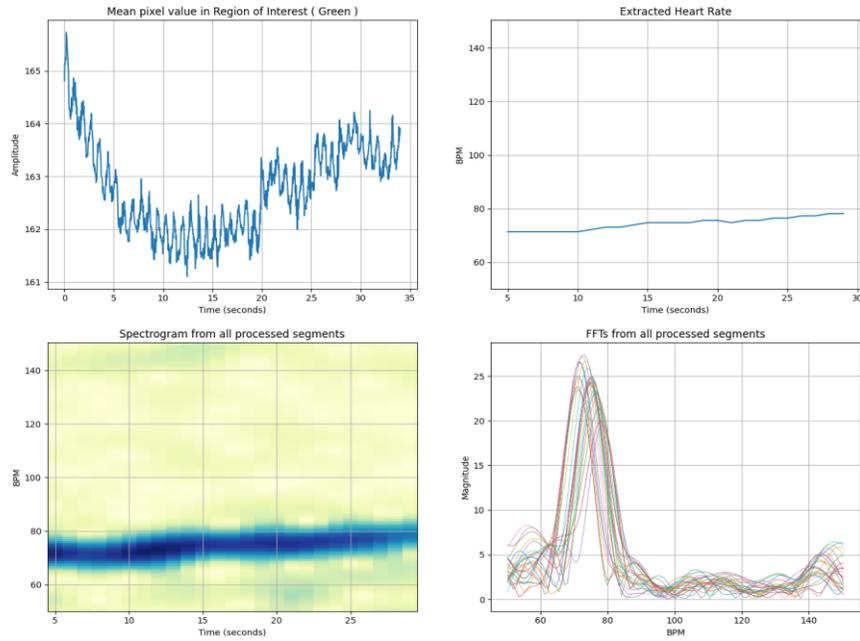


Figure 49. The VM algorithm results for an iPhone 12 Pro Max video with smart Auto 30FPS turned on

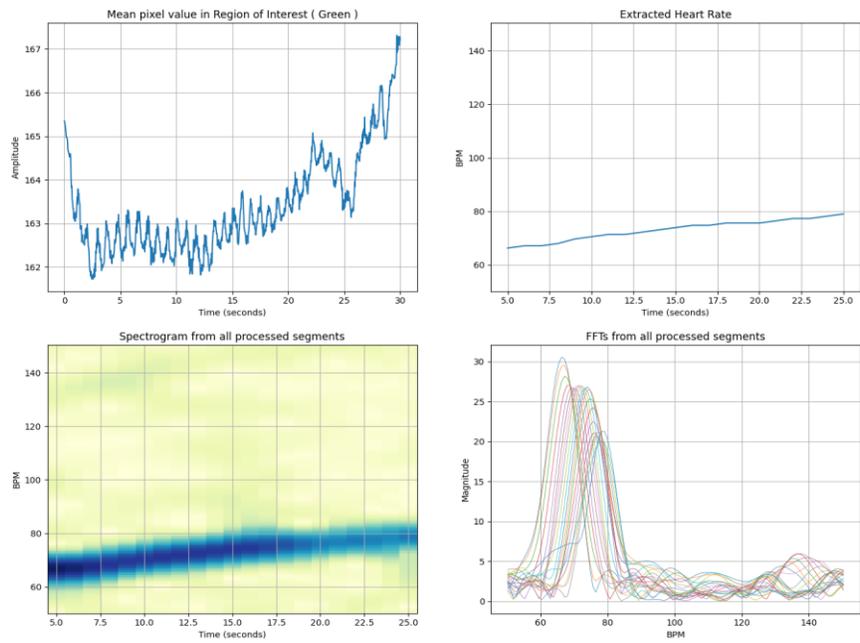


Figure 50. The VM algorithm results for an iPhone 12 Pro Max video with smart Auto 30&60 FPS turned on

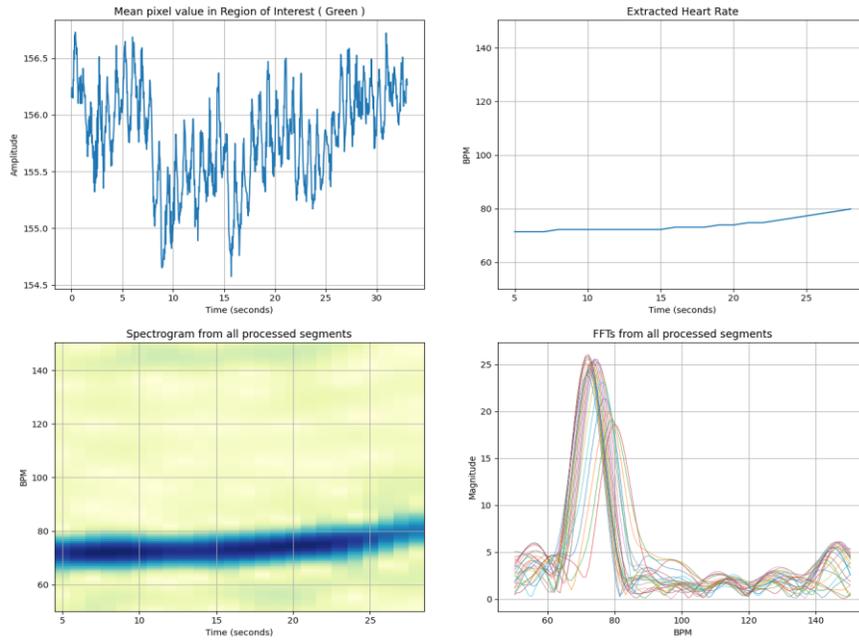


Figure 51. The VM algorithm results for an iPhone 12 Pro Max video with smart Auto 30&60FPS and HDR turned on

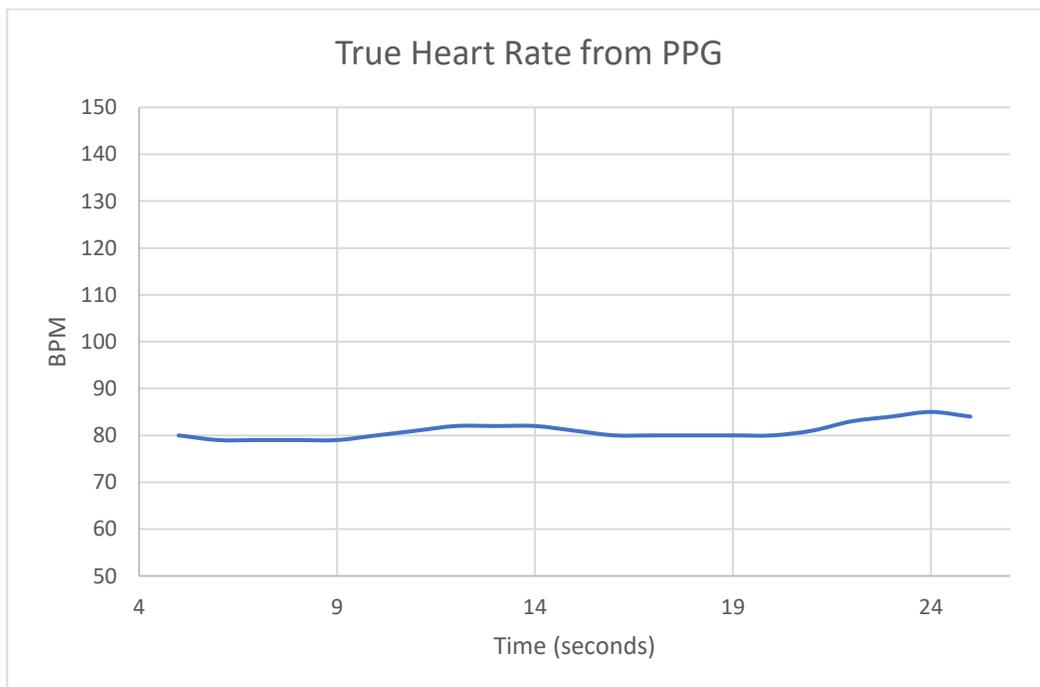


Figure 52. True heart rate results from PPG for iPhone 12 Pro Max videos

Next, the effect of noise from different illumination levels resulting from two different lighting conditions is discussed.

5.3 Effect of Illumination Level on the Video Magnification Algorithm

This section's contents are based on the paper in review for publication at the 2022 IEEE Sensors Applications Symposium conference titled *Effects of Lighting and Window Length on Heart Rate Assessment through Video Magnification* [11]

Different light sources had been used in the literature with Video Magnification techniques for heart rate assessment. Some works have neglected to mention the light source used. Out of the mentioned light sources, some works have used ambient light [30], [34], [36], [47], [59], some have used fluorescent and incandescent light [29] and others have used LED light [51] including our research team [10]. Nevertheless, no work has looked at examining the effect of the source spectrum or the illumination level on the Video Magnification algorithm for heart rate detection.

This section compares the effect of two light sources on the performance of the Video Magnification algorithm for 19 subjects under two illumination levels and light sources. The first illumination source is LED only while the second illumination combines LED with Halogen. The videos used to generate the results of this chapter are the ones obtained from the NRC group.

5.3.1 Analysis Details

Two NRC test conditions C1 and C2 from Table 1 which both correspond to the participants in seated position with mask off but under two different lighting configurations are compared for 19 subjects. In C1, the lighting configuration was LED only (which is the condition that has been used in all sections of this thesis thus far that used NRC video). In C2, the lighting configuration was LED combined with Halogen light. Table 7 shows a summary of the two conditions.

Table 7. Summary of the two conditions used in the analysis

Condition:	Position and Mask	Lighting Configuration
C1	Sitting & Mask Off	LED only
C2		LED + Halogen

The illumination levels (light energy per wavelength) of the two lighting configurations were measured using a Mettue SIM-2 Spectral Irradiance Meter and the results can be seen in Figure 53 plotted against the light wavelengths. The focus of the VM algorithm analysis is on the green band since it had been shown to have the best signal to noise ratio for heart rate from RGB video [60]. Therefore, the illumination levels of focus in this chapter lie over the wavelengths of 480-550nm and this area is identified with dotted green lines in Figure 53.

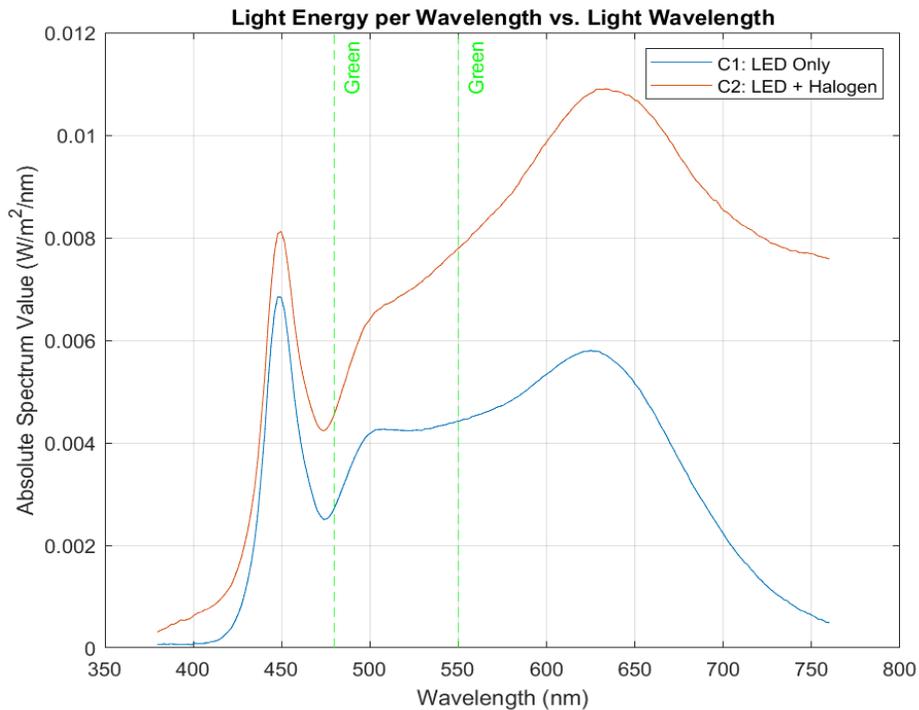


Figure 53. Illumination levels of the two lighting conditions C1 and C2 obtained from [11]

Condition C2 that combined Halogen light with LED has a much higher energy per wavelength

for the green colour spectrum over condition C1 with LED only. 38 videos in total were analyzed using the Video Magnification algorithm for these two lighting configurations. The Video Magnification algorithm was configured with a window segment length of 10 s with 9 s overlap between adjacent windows. It was also configured to span a bandpass filter range of 50-150 BPM for the full duration of the 90 s videos for the 19 subjects. The FFT size was set to 2048.

Heart rate results from the green channel of the VM algorithm were then compared to the true heart rate results of the participants obtained from the ground truth PPG using the % mean absolute error metric in Equation 4.

5.3.2 Effect of Two Illumination Levels from the Combination of Two Light Sources

The individual % mean absolute errors for the 19 subjects is summarized in Table 8 for the two lighting configurations C1 and C2. Table 9 contains the average, standard deviation and maximum of the two columns in Table 8.

Table 8. % mean absolute error results for the two lighting configurations from [11]

Subject ID	% Mean Absolute Error	
	C1: LED Only	C2: LED + Halogen
S1	3.1	4.8
S2	1.5	2.1
S3	0.7	0.8
S4	11.6	2.1
S5	3.4	0.9
S6	4.2	6.6
S7	1.6	1.0
S8	4.6	0.8
S9	2.2	1.4
S10	1.2	2.4
S11	20.5	18.1
S12	8.0	1.2
S13	5.7	2.4
S14	0.5	0.7
S15	0.6	1.2
S16	3.9	6.6
S17	1.5	1.4
S18	20.7	16.6
S19	13.6	3.7

Table 9. Average, standard deviation and maximum % mean absolute error for the 19 subjects from [11]

Testing Condition	C1: LED only	C2: LED + Halogen
Average	5.7	3.9
Standard Dev.	6.2	4.9
Maximum	20.7	18.1

From Table 9 it was found that C2 has a lower average and standard deviation of the % mean absolute error for the 19 subjects over C1 ($3.0 \pm 4.9\%$ vs. $5.7 \pm 6.2\%$ respectively). This improvement in the Video Magnification performance with the addition of Halogen light can be explained by the increase in illumination level that this light provided when it was combined with LED. Nevertheless, it is important to note that the videos for the two conditions of the same subject were taken subsequently and not at the same time, and therefore motion present in the two videos might have varied between the two videos.

For subject S4 the % mean absolute error for C1 was 11.6% and it was 2.1% in C2. Aside from the change in light source, the subject was moving significantly in the C1 video (yawning), and they were not moving as much in the C2 video. The motion from yawning caused the ROI to be missing from some of the video frames, and that affected the results significantly. The VM algorithm results for S4 can be seen in Figure 54 for condition C1 and Figure 55 for condition C2 with the true heart rate in Figure 56.

The maximum % mean absolute error for C1 according to Table 9 of 20.7% corresponds to subject S18 in Table 8. Subject S18 had the darkest skin tone amongst the 19 subjects, therefore this poor performance of the Video Magnification is attributed to its sensitivity with darker skin tones [20]. S18 also had a high % mean absolute error of 16.6% for C2.

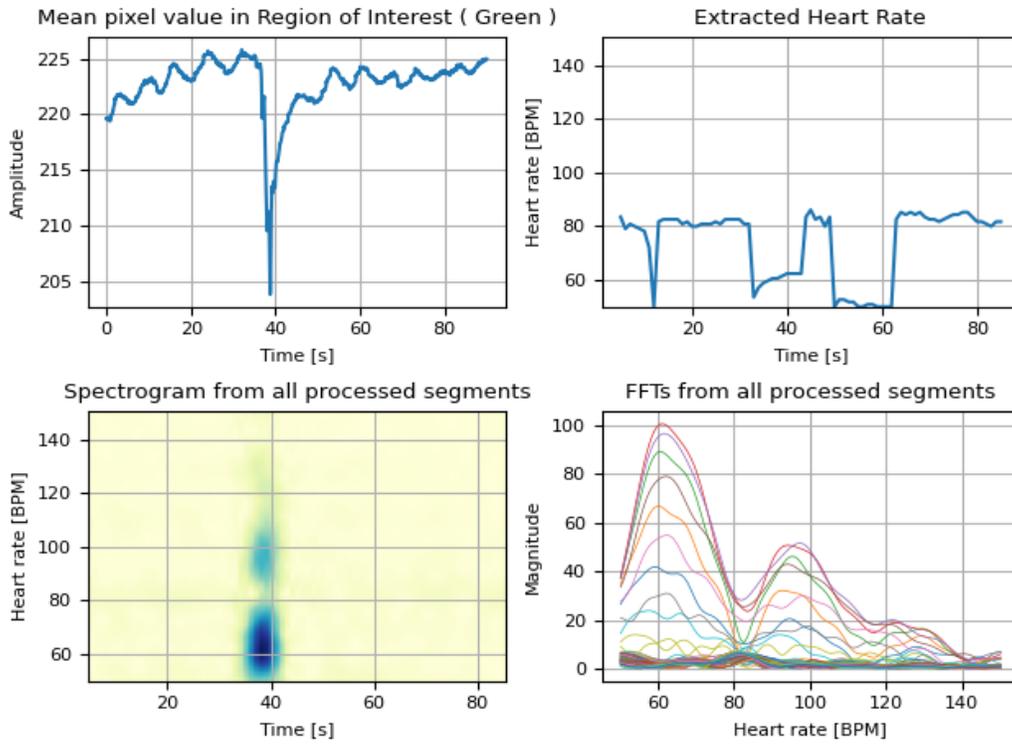


Figure 54. The VM algorithm results for condition C1 subject S4 (yawning)

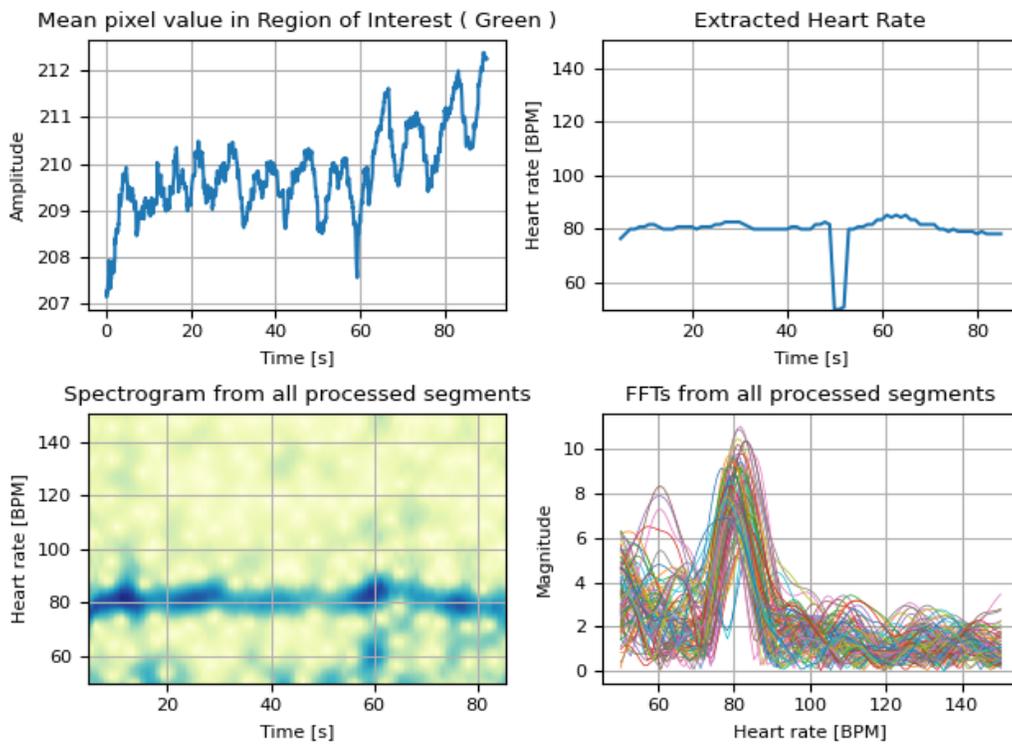


Figure 55. The VM algorithm results for condition C2 subject S4

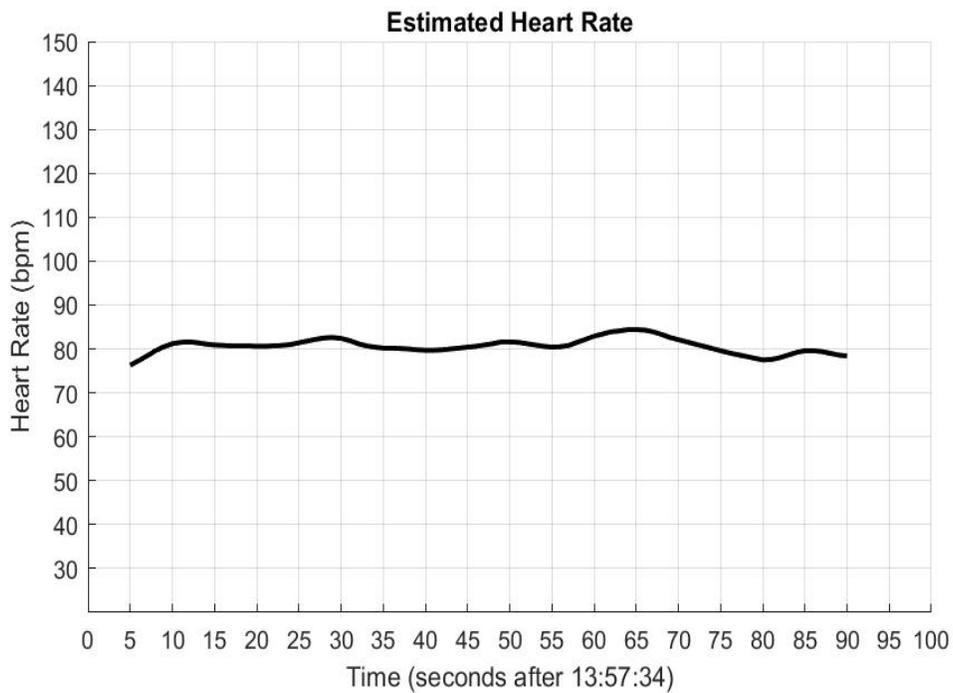


Figure 56. True Heart Rate Results from PPG Subject S4

The maximum % mean absolute error for C2 according to Table 9 of 18.1% corresponds to subject S11 in Table 8. This subject had a fair skin tone but experienced significant head motion in the two videos of C1 and C2. The subject also had a % mean absolute error of 20.5% for C1. The VM algorithm results for subject S11 condition C2 can be seen in Figure 57 and their true heart rate can be seen in Figure 58. The spectrogram plot in Figure 57 contains the subject's true heart rate between 80 and 100 BPM seen in Figure 58 but it also contains other frequency components that are affecting the extracted heart rate results of the VM algorithm. The extracted heart rate from the VM algorithm is what is compared to the true HR results from the PPG and therefore, a large % mean absolute error is resulting.

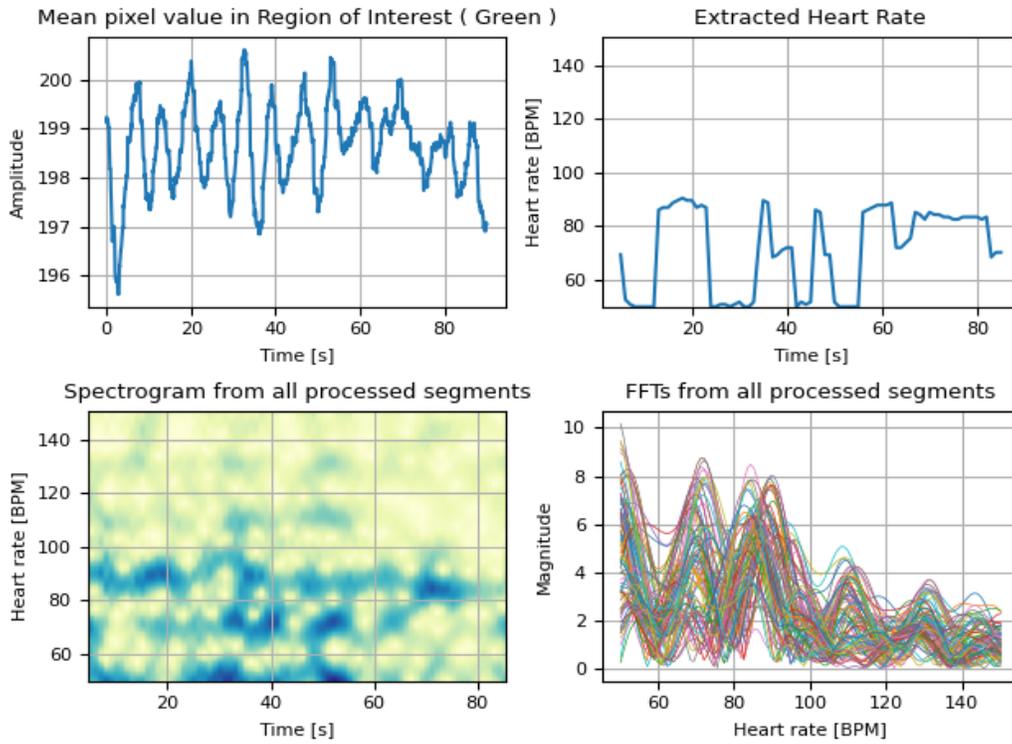


Figure 57. VM algorithm results for subject S11 condition C2 obtained from [11]

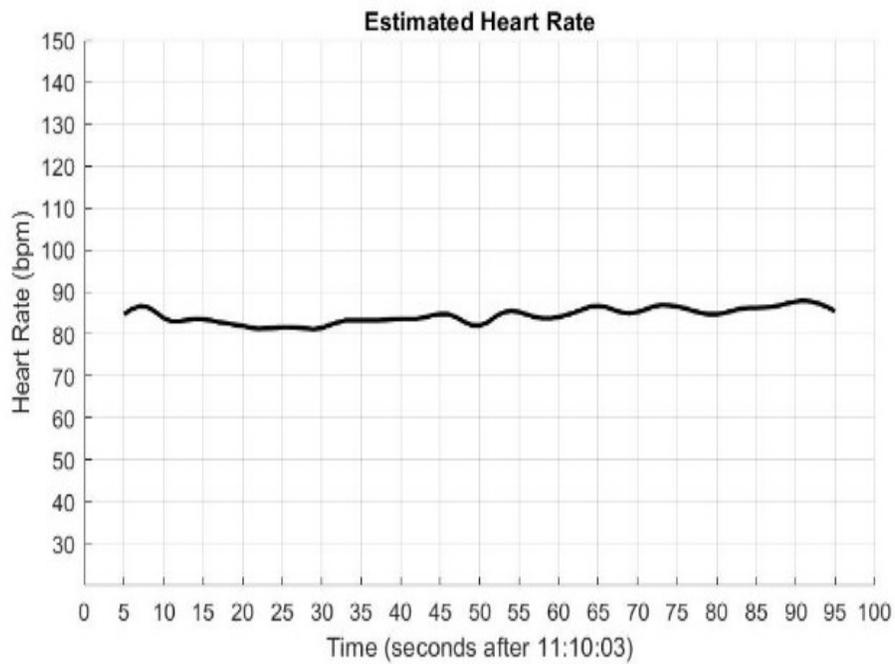


Figure 58. True heart rate from PPG for subject S11 condition C2 obtained from [11]

Chapter 6: Conclusions

This chapter contains the conclusions of the thesis contributions followed by potential future work.

6.1 Conclusion of Thesis Contributions

One contribution of the thesis explored the effect of the size of the ROI on the performance of the Video Magnification algorithm. It was found that the detection of heart rate is possible from very small ROIs when multiple ROIs were averaged together. It was also found that increasing the ROI size resulted in improved accuracy of heart rate detection by the VM algorithm when the background was not included. Detecting heart rate from small ROIs would be beneficial if a large area is not available in the video due to occlusion of hair or some other factor. The effect of motion was found to have detrimental effects on the accuracy of the VM algorithm when the subject moved from sitting to standing. This suggested that a form of motion tracking could be investigated to determine if it could help reduce the error. Lastly, it was found that averaging the pixel values within a larger ROI to detect heart rate was superior to detecting the heart rate from multiple smaller adjacent windows and then averaging them as shown in Table 10 below.

Table 10. Comparison of the percent mean absolute error for the four conditions (C1-C4) using small and large ROI from [10]

Condition	Small ROI: 16x16	Large ROI: 64x64
C1	9.9%	4.4%
C2	9.9%	4.5%
C3	28.9%	16.6%
C4	27.0%	18.3%

Another contribution of this thesis explored the effect of increasing the window segment length from 10s to 20s to 30s on the Video Magnification algorithm. Two conclusions were drawn related

to the subject's head motion in the video. In the first one it was found that increasing the window segment length is advantageous when the subject displays significant head motion in the video. In the second one, it was found that when the head motion is minimal in the video, increasing the window segment length above 10s reduced the performance of the Video Magnification algorithm at detecting small momentary changes in heart rate.

Another contribution of this thesis investigated the performance of the Video Magnification algorithm in detecting heart rate from artificial videos using large and small signal amplitudes without noise and small signal amplitudes with noise. The Video Magnification algorithm correctly detected the true heart rates for all videos containing the large signal amplitudes and the amplitude of the FFT was found to be linearly correlated with the signal amplitude. Moreover, the Video Magnification algorithm correctly detected the true heart rate for the videos containing the small signal amplitudes down to a signal amplitude of 0.002 in the absence of noise. Additionally, in the presence of noise, the Video Magnification algorithm's accuracy at detecting the true heart rate of a signal amplitude of 0.002 improved, which is likely due to the dithering effect of the noise frequency and the signal harmonics. Lastly, at high noise levels ($m=25$), the effect on the small signal amplitudes (0.005-0.003) was very similar although the 0.005 amplitude is 66% larger than the 0.003 amplitude.

Moreover, another thesis contribution investigated the effect of smart video features of modern smartphone camera systems on the performance of the Video Magnification algorithm in detecting heart rate from facial video. It was found that the VM algorithm's ability to detect subject heart rates from videos recorded using a basic VGA webcam is not affected with the iPhone 12 Pro Max's high-quality lens in the absence of smart software features. However, it was found that the smart software features that improve the captured video aesthetically have a detrimental effect on

the VM algorithm's ability to detect heart rate.

The last contribution of this thesis investigated the effect of increasing the illumination level on the performance of the Video Magnification algorithm in detecting heart rate. The two illumination levels were provided by LED only (lower illumination level) and Halogen combined with LED (higher illumination level). It was found that with the addition of Halogen and the increase in illumination that it provided, the Video Magnification algorithm results improved on average for the 19 subjects tested of varying skin tones, age and gender.

6.2 Guidelines for Video Capture and Optimal Video Magnification Results

Based on the findings of the work summarized in this thesis, the video capture process that could result in ideal results from the VM algorithm include:

- The captured subject being in sitting position with their head resting on a head rest and their arms resting on arm rests while facing the camera directly to reduce involuntary motion.
- The optimal light source would be LED combined with Halogen light placed on either side between the subject and the camera and illuminating the face.
- The camera would be placed on a tripod to eliminate any motion.
- The ROI to be processed would be centered on the forehead and contains the entire forehead.
- Lastly, the window segment length would be set to 30 seconds.

6.3 Suggestions for Future work

In the future, the simulated signals described in **Chapter 5 section 5.1: Effect of Quantization Noise on the Video Magnification Algorithm**, could be compared to real signals containing noise from real sources in order to further understand the performance of the VM algorithm.

Additionally, the effect of varying the intensity of one light source (ex. LED) on the performance of the VM algorithm could be explored in order to distinguish between the effect of multiple light sources causing higher illumination vs. higher illumination levels provided by the same light source.

References

- [1] E. Mejía-Mejía, J. M. May, R. Torres, and P. A. Kyriacou, “Pulse rate variability in cardiovascular health: a review on its applications and relationship with heart rate variability,” *Physiol. Meas.*, vol. 41, no. 7, p. 07TR01, Aug. 2020.
- [2] M.-H. Cheng, L.-C. Chen, Y.-C. Hung, C. M. Yang, and T. L. Yang, “A Real-Time Heart-Rate Estimator from Steel Textile ECG Sensors in a Wireless Vital Wearing System,” in *2008 2nd International Conference on Bioinformatics and Biomedical Engineering*, Shanghai, China, May 2008, pp. 1339–1342.
- [3] D. Biswas, N. Simoes-Capela, C. Van Hoof, and N. Van Helleputte, “Heart Rate Estimation From Wrist-Worn Photoplethysmography: A Review,” *IEEE Sens. J.*, vol. 19, no. 16, pp. 6560–6570, Aug. 2019.
- [4] B. Wallace, L. Y. Kassab, A. Law, R. Goubran, and F. Knoefel, “Contactless Remote Assessment of Heart Rate and Respiration Rate Using Video Magnification,” *IEEE Instrum. Meas. Mag.*, vol. 25, no. 1, pp. 20–27, Feb. 2022.
- [5] 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,” *ACM Trans. Graph. TOG*, vol. 31, no. 4, p. 8, 2012.
- [6] G. de Haan and V. Jeanne, “Robust Pulse Rate From Chrominance-Based rPPG,” *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2878–2886, Oct. 2013.
- [7] S. Himmel, M. Ziefle, and K. Arning, “From Living Space to Urban Quarter: Acceptance of ICT Monitoring Solutions in an Ageing Society,” in *Human-Computer Interaction. Users and Contexts of Use*, Berlin, Heidelberg, 2013, pp. 49–58.
- [8] M. A. Hassan, A. S. Malik, D. Fofi, B. Karasfi, and F. Meriaudeau, “Towards health monitoring using remote heart rate measurement using digital camera: A feasibility study,” *Measurement*, vol. 149, p. 106804, Jan. 2020.
- [9] L. Yassin Kassab, B. Wallace, R. Goubran, and F. Knoefel, “Video Magnification Remote Sensing of Heart Rate and Respiration Rate,” report produced for NRC Canada, 2021.
- [10] L. Yassin Kassab, A. Law, B. Wallace, J. Larivière-Chartier, R. Goubran, and F. Knoefel, “Effects of Region of Interest Size on Heart Rate Assessment through Video Magnification,” in *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2021, p. 6.
- [11] L. Yassin Kassab, A. Law, B. Wallace, R. Goubran, F. Knoefel, and J. Larivière-Chartier, “Effects of Lighting and Window Length on Heart Rate Assessment through Video Magnification,” in review for 2022 Sensors Applications Symposium, 2022, p. 6.
- [12] L. Y. Kassab, A. Law, B. Wallace, J. Lariviere-Chartier, R. Goubran, and F. Knoefel, “The Effect of Noise on Contactless Heart Rate Measurement using Video Magnification,” in *accepted at 2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, May 2022, p. 6.
- [13] M. A. Hassan *et al.*, “Heart rate estimation using facial video: A review,” *Biomed. Signal Process. Control*, vol. 38, pp. 346–360, Sep. 2017.
- [14] S. Leonhardt, L. Leicht, and D. Teichmann, “Unobtrusive Vital Sign Monitoring in Automotive Environments—A Review,” *Sensors*, vol. 18, no. 9, p. 3080, Sep. 2018.
- [15] A. M. Unakafov, “Pulse rate estimation using imaging photoplethysmography: generic framework and comparison of methods on a publicly available dataset,” *Biomed. Phys. Eng. Express*, vol. 4, no. 4, p. 045001, Apr. 2018.

- [16] F. Khanam, A. Al-Naji, and J. Chahl, "Remote Monitoring of Vital Signs in Diverse Non-Clinical and Clinical Scenarios Using Computer Vision Systems: A Review," *Appl. Sci.*, vol. 9, no. 20, p. 4474, Oct. 2019.
- [17] X. Chen, J. Cheng, R. Song, Y. Liu, R. Ward, and Z. J. Wang, "Video-Based Heart Rate Measurement: Recent Advances and Future Prospects," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 3600–3615, Oct. 2019.
- [18] M. Harford, J. Catherall, S. Gerry, J. Young, and P. Watkinson, "Availability and performance of image-based, non-contact methods of monitoring heart rate, blood pressure, respiratory rate, and oxygen saturation: a systematic review," *Physiol. Meas.*, vol. 40, no. 6, p. 06TR01, Jul. 2019.
- [19] C. H. Antink, S. Lyra, M. Paul, X. Yu, and S. Leonhardt, "A Broader Look: Camera-Based Vital Sign Estimation across the Spectrum," *Yearb. Med. Inform.*, vol. 28, no. 01, pp. 102–114, Aug. 2019.
- [20] Y. Deng and A. Kumar, "Standoff heart rate estimation from video – a review," in *Mobile Multimedia/Image Processing, Security, and Applications 2020*, Online Only, United States, Apr. 2020, p. 6.
- [21] H. I. Shahadi, H. J. Albattat, Z. J. Al-allaq, and A. T. Thahab, "Eulerian video magnification: a review," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 18, no. 2, p. 799, May 2020.
- [22] X. He, "Signal Enhancement Applied to Pulse Transit Time Measurement," Doctor of Philosophy, Carleton University, Ottawa, Ontario, 2016.
- [23] S. Bennett, "Non-Contact Bed-Based Monitoring of Vital Signs," Doctor of Philosophy, Carleton University, Ottawa, Ontario, 2019.
- [24] M. Hu *et al.*, "Combination of near-infrared and thermal imaging techniques for the remote and simultaneous measurements of breathing and heart rates under sleep situation," *PLOS ONE*, vol. 13, no. 1, p. e0190466, Jan. 2018.
- [25] S. L. Bennett, R. Goubran, and F. Knoefel, "Comparison of motion-based analysis to thermal-based analysis of thermal video in the extraction of respiration patterns," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo, Jul. 2017, pp. 3835–3839.
- [26] G. Scebba, G. Da Poian, and W. Karlen, "Multispectral Video Fusion for Non-Contact Monitoring of Respiratory Rate and Apnea," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 1, pp. 350–359, Jan. 2021.
- [27] J. Nikolic-Popovic and R. Goubran, "Impact of motion artifacts on video-based non-intrusive heart rate measurement," in *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Benevento, Italy, May 2016, pp. 1–6.
- [28] A. Alzahrani and A. Whitehead, "Preprocessing Realistic Video for Contactless Heart Rate Monitoring Using Video Magnification," in *2015 12th Conference on Computer and Robot Vision*, Halifax, NS, Canada, Jun. 2015, pp. 261–268.
- [29] Y. S. Dosso, A. Bekele, and J. R. Green, "Eulerian Magnification of Multi-Modal RGB-D Video for Heart Rate Estimation," in *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Rome, Jun. 2018, pp. 1–6.
- [30] X. He, R. A. Goubran, and X. P. Liu, "Wrist pulse measurement and analysis using Eulerian video magnification," in *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, Las Vegas, NV, USA, Feb. 2016, pp. 41–44.

- [31] X. He *et al.*, “Video-Based Analysis of Heart Rate Applied to Falls,” in *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Rome, Italy, Jun. 2018, pp. 1–5.
- [32] C. Gonzalez Viejo, S. Fuentes, D. Torrico, and F. Dunshea, “Non-Contact Heart Rate and Blood Pressure Estimations from Video Analysis and Machine Learning Modelling Applied to Food Sensory Responses: A Case Study for Chocolate,” *Sensors*, vol. 18, no. 6, p. 1802, Jun. 2018.
- [33] H. Lauridsen *et al.*, “Extracting physiological information in experimental biology via Eulerian video magnification,” *BMC Biol.*, vol. 17, no. 1, p. 103, Dec. 2019.
- [34] X. He, R. A. Goubran, and X. P. Liu, “Using Eulerian video magnification framework to measure pulse transit time,” in *2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Lisboa, Portugal, Jun. 2014, pp. 1–4.
- [35] M.-Z. Poh, D. J. McDuff, and R. W. Picard, “Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.
- [36] S. L. Bennett, R. Goubran, and F. Knoefel, “Adaptive eulerian video magnification methods to extract heart rate from thermal video,” in *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Benevento, Italy, May 2016, pp. 1–5.
- [37] S. L. Bennett, R. Goubran, and F. Knoefel, “Examining the Effect of Noise on Biosignal Estimates Extracted through Spatio-Temporal Video Processing,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, Jul. 2019, pp. 4504–4508.
- [38] S. Bennett, T. N. El Harake, R. Goubran, and F. Knoefel, “Adaptive Eulerian Video Processing of Thermal Video: An Experimental Analysis,” *IEEE Trans. Instrum. Meas.*, vol. 66, no. 10, pp. 2516–2524, Oct. 2017.
- [39] S. L. Bennett, R. Goubran, and F. Knoefel, “The detection of breathing behavior using Eulerian-enhanced thermal video,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Aug. 2015, pp. 7474–7477.
- [40] M. van Gastel, S. Stuijk, and G. de Haan, “Motion Robust Remote-PPG in Infrared,” *IEEE Trans. Biomed. Eng.*, vol. 62, no. 5, pp. 1425–1433, May 2015.
- [41] X. He, R. Goubran, and F. Knoefel, “IR night vision video-based estimation of heart and respiration rates,” in *2017 IEEE Sensors Applications Symposium (SAS)*, Glassboro, NJ, USA, 2017, pp. 1–5.
- [42] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, “Algorithmic Principles of Remote PPG,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1479–1491, Jul. 2017.
- [43] W. Wang, A. C. den Brinker, and G. de Haan, “Full video pulse extraction,” *Biomed. Opt. Express*, vol. 9, no. 8, p. 3898, Aug. 2018.
- [44] H. E. Tasli, A. Gudi, and M. den Uyl, “Remote PPG based vital sign measurement using adaptive facial regions,” in *2014 IEEE International Conference on Image Processing (ICIP)*, Paris, France, Oct. 2014, pp. 1410–1414.
- [45] M.-H. Hu, G.-T. Zhai, D. Li, Y.-Z. Fan, X.-H. Chen, and X.-K. Yang, “Synergetic use of thermal and visible imaging techniques for contactless and unobtrusive breathing measurement,” *J. Biomed. Opt.*, vol. 22, no. 03, p. 1, Mar. 2017.

- [46] Y. Qiu, Y. Liu, J. Arteaga-Falconi, H. Dong, and A. El Saddik, "EVM-CNN: Real-Time Contactless Heart Rate Estimation From Facial Video," *IEEE Trans. Multimed.*, vol. 21, no. 7, pp. 1778–1787, Jul. 2019.
- [47] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Express*, vol. 16, no. 26, p. 21434, Dec. 2008.
- [48] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Express*, vol. 18, no. 10, p. 10762, May 2010.
- [49] L. Liu, L. Lu, J. Luo, J. Zhang, and X. Chen, "Enhanced Eulerian video magnification," in *2014 7th International Congress on Image and Signal Processing*, Dalian, China, Oct. 2014, pp. 50–54.
- [50] X. He, "Signal Enhancement Applied to Pulse Transit Time Measurement," Doctor of Philosophy, Carleton University, Ottawa, Ontario, 2016.
- [51] F. P. Wieringa, F. Mastik, and A. F. W. van der Steen, "Contactless Multiple Wavelength Photoplethysmographic Imaging: A First Step Toward 'SpO2 Camera' Technology," *Ann. Biomed. Eng.*, vol. 33, no. 8, pp. 1034–1041, Aug. 2005.
- [52] M. Mozafari, A. Law, S. Béni Tchoudem Djouaka, J. Green, and R. Goubran, "Blind Source Separation for Respiration Rate from Depth Camera," accepted at 2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2022.
- [53] M. Mozafari, A. Law, J. Green, and R. Goubran, "Respiration Rate from Thermal Camera Using Tensor Decomposition," accepted at 2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2022.
- [54] A. Law, "Multimodal contactless physiological measurement for health screening applications REB Protocol." NRC, 2020.
- [55] L. I. Castillo, M. E. Browne, T. Hadjistavropoulos, K. M. Prkachin, and R. Goubran, "Automated vs. manual pain coding and heart rate estimations based on videos of older adults with and without dementia," *J. Rehabil. Assist. Technol. Eng.*, vol. 7, Jan. 2020.
- [56] L. F. Corral Martinez, G. Paez, and M. Strojnik, "Optimal wavelength selection for noncontact reflection photoplethysmography," Puebla, Mexico, Aug. 2011, p. 801191.
- [57] E. B. Blackford, J. R. Estep, and D. J. McDuff, "Remote spectral measurements of the blood volume pulse with applications for imaging photoplethysmography," in *Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics*, San Francisco, United States, Feb. 2018, p. 41.
- [58] Y. Qiu, "Real-Time Contactless Heart Rate Estimation from Facial Video," PhD diss, Université d'Ottawa/University of Ottawa, 2018.
- [59] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. Clifton, and C. Pugh, "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models," *Physiol. Meas.*, vol. 35, no. 5, pp. 807–831, May 2014.
- [60] S. Zaunseder, A. Trumpp, D. Wedekind, and H. Malberg, "Cardiovascular assessment by imaging photoplethysmography – a review," *Biomed. Eng. Biomed. Tech.*, vol. 63, no. 5, pp. 617–634, Oct. 2018.