HEART ACTIVITY MONITOR CAMERA: HEART RATE ACQUISITION FROM VIDEO

Audrey Chung **Jan Kulinski** **Joanne Leong**

Professor Alex Wong

Department of Systems Design Engineering

Abstract: This paper provides an overview of the team's project to create a video-based system for acquiring the heart rate of multiple subjects. The system combines facial and feature detection, point tracking, and signal processing techniques to isolate a heart rate for multiple subjects in a given video. Implementation details are provided, and associated challenges and weaknesses are addressed. The implications of the work are discussed and potential future work is outlined to improve the accuracy and utility of the system.

1. INTRODUCTION

Technological advancements in the field of medicine have consistently increased the accuracy and lowered the cost of taking biometric measurements. As a result, medical diagnosis and health tracking has become easier and more financially viable over time. Using a process called Eulerian Video Magnification (EVM), a team of researchers at MIT have shown that it is possible to isolate a person's heart rate from a video of still, exposed skin by amplifying the colour change of skin flushing when blood flows through it [1].

Heart rate is a useful metric for assessing the condition of a person's body. It is simple to measure, even without hardware, by counting the number of beats felt over a period of time. However, a video-based system offers the potential benefit of continuous, non-invasive monitoring.

A small number of mobile applications currently on the market attempt to use phone cameras to read a person's heart rate. In general, they are slow and provide unreliable readings. This is largely due to the technical challenges involved in extracting a pulse from the complex input signal of video, discussed in more detail in section 5: technical challenges.

2. PROBLEM STATEMENT

The Heart Activity Monitor Camera (HAM Cam) aims to provide continuous heart rate monitoring from video for multiple subjects. In hospital applications, this will allow for non-invasive vitals monitoring of multiple patients in a given area without the need to connect additional equipment for each subject.

3. GOALS

The project's primary goal is to acquire an accurate estimate of a person's heart rate from video. In order for the system to be feasible for real world applications, the following were outlined as constraints on the system's functions:

1. Require no more than 10 seconds of video to generate a reading
2. An error tolerance of at most 3 bpm
3. Ability to read a heartbeat for multiple subjects from a single video
4. Track the subjects' heart rates over time
5. Detect abnormalities and generate alerts

4. SOLUTION

The solution was developed as a post-processing technique, working with pre-recorded videos of finite length rather than a feed of video. The processing is performed in MATLAB which provides ample tools for the image and signal processing techniques used.

The EVM process amplifies any changes in colour or position. While this is able to bring out skin colour changes over time, any motion of the subject is also amplified. The amplification of subject motion tends to dominate the processed video, creating a mess of amplified motions from which skin colour changes are not recoverable.

In order to counteract the effects of motion amplification dominating the processed video, the solution uses point tracking to amplify only local colour changes. To start tracking a new subject, they are identified in the video by running a face detection algorithm [2] on individual frames. From the identified face, the eyes are detected in order to place points to track on their forehead and cheeks. These points tend to flush with blood flow and avoid interference from shadows and hairline.

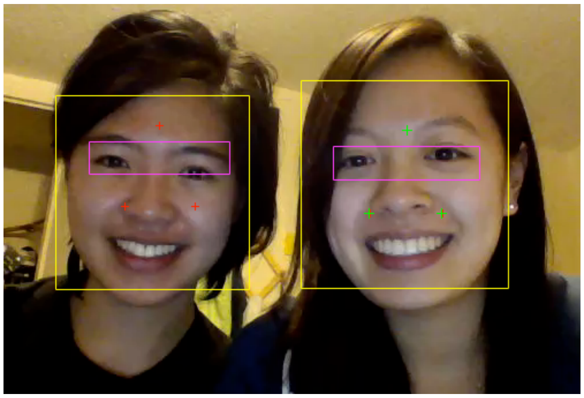

Figure 1: Identified faces, eye pairs and points to track

In the following frames of video, the three points on each face are updated so that they remain in the same location on the subjects' faces. The solution creates a sub-video around each tracked point by cropping a small area (11x11 pixels). Each of the three sub-videos is then processed using EVM, and signals for the three colour components (RGB) are generated using the average intensity of the frame.

The nine signals are then processed using Independent Component Analysis (ICA) [3]. This process decomposes the signals into independent components. This results in nine signals that represent frequencies present in the RGB values. The underlying theory is that the heart beat is present as one of the contributing signals and may be isolated by decoupling it from the noise.

The majority of these resultant signals are very low frequency, corresponding to DC noise. All frequencies outside of the range of reasonable human heart rates (35 BPM to 180 BPM, or ~0.6Hz to 3Hz) are eliminated as potential heart rates. This typically narrows the search to one or two resultant frequencies. Of the remaining signals, the one with the strongest response is chosen and returned as the heart rate for the subject.

The next part of the solution would be to take continuous measurements for each subject, and providing a graphical user interface (GUI) to monitor the heart rate over time. Due to time constraints, the algorithmic MATLAB component and the constructed GUI were not integrated, and are not able to function together autonomously.

5. Technical Challenges

Several design aspects may be detrimental to the quality of the final reading. Low or changing light conditions increase the difficulty of identifying faces and features, as well as track points in the video. The camera used may also be a source of noise, decreasing the video's signal to noise ratio and making it more difficult to isolate the true heart rate. Furthermore, losing one or more of the tracked points could make the final reading less robust to noise, thereby compounding the issue.

6. Future Work

The system was able to achieve an average error of 7.25% ± 7.2457%; however, there is no way to automatically determine the accuracy per video. Tracking additional points, using stochastic sampling, using multiple runs and parameters, and integrating intelligent techniques to determine optimal parameters may improve the accuracy.

7. Discussion and Conclusions

The final system is not yet robust enough for the target application of monitoring the vitals of multiple subjects in a hospital context, where the accuracy and real-time aspects of heart rate readings are of high importance. However, the acquisition of heart rate from video may be used outside of the medical field where it's not as mission critical (for instance, it can be used for a personal health tracking and analytics system).

A system capable of reading heart rate from video brings up questions regarding privacy. One's heart rate could be considered as personal data which would be at greater risk of being exposed. Depending on the application and the context, people may have different levels of sensitivity about this, for instance being more willing to share this information in a hospital waiting room as opposed to an interrogation room. This technology will continue to develop and begin to be implemented in the world, at which point society as a whole will have to answer the question of how to regulate biometric data.

8. References

[1] Wu, H. et al., *Eularian Video Magnification for Revealing Subtle Changes in the World*, ACM Transactions on Graphics, Vol. 31, No. 4, 2012

[2] Viola, Jones., *Rapid object detection using a boosted cascade of simple features,* Vision and Pattern Recognition, 2001.

[3] Poh et al. *Advancements in Noncontact, Muliparameter Phusiological Measurements Using a Webcam*. IEE Transactions on Biomedical Engineering, Vol. 58, No. 1, Jan 2011.