

Measuring Heart Rate from Video

Isabel Bush

ibush@stanford.edu

Stanford Computer Science

353 Serra Mall, Stanford, CA 94305

Introduction

A person's heart rate can be indicative of their health, fitness, activity level, stress, and much more. Cardiac pulse is typically measured in clinical settings using electrocardiogram (ECG), which requires patients to wear chest straps with adhesive gel patches that can be abrasive and become uncomfortable for the user. Chest straps are also common for fitness heart-rate monitors used by joggers or other athletes. Heart rate may also be monitored using pulse oximetry sensors that may be worn on the fingertip or earlobe. These sensors are not convenient for long-term wear and the pressure can become uncomfortable over time.

In addition to the discomforts of traditional pulse measurement devices, these devices can damage the fragile skin of premature newborns or elderly people. For these populations especially, a non-contact means of detecting pulse could be very beneficial. Non-contact heart rate measurement through a simple webcam would also aid telemedicine and allow the average person to track their heart rate without purchasing special equipment. As part of the recent gain in popularity of fitness apps and the quantified self, regular non-obtrusive monitoring through a computer web camera may help detect changes in a person's heart rate over time and indicate changing health or fitness.

Heart rate can be detected without contact through photo-plethysmography (PPG), which

measures variations in blood volume by detecting changes in light reflectance or transmission throughout the cardiovascular pulse cycle. PPG is usually performed with dedicated light sources with red or infrared wavelengths, as is the case for pulse oximetry sensors.

Verkruysse *et al.* showed that the plethysmographic signal could also be detected in video from a regular color camera [5]. They found that the signal could be detected within the red, green, and blue channels of color video of exposed skin, but that it was strongest in the green channel, which corresponds to the fact that hemoglobin has absorption peaks for green and yellow light wavelengths. They also found that although the signal could be detected in multiple locations on the body, it was strongest on the face, especially on the forehead.

Although the plethysmographic signal may be detected in the raw color channel data, it is mixed in with other sources of color variation such as changes in ambient light or motion. Poh *et al.* found that the signal could be better extracted by using independent component analysis (ICA) to separate independent source signals from the mixed color signals [4].

Other studies have shown that color changes in the face due to pulse may be magnified by amplifying small changes between video frames [7], and that heart rate can be detected through vertical head motion in addition to color changes [1]. Although these are interesting new developments

in this space, they are less practical for daily or medical use as the former is more for visualization than quantification, and the latter requires the subject to remain very still for accurate measurements.

In this project, I will explore an approach for heart rate detection using RGB color changes in video of faces similar to that done by Poh *et al.* [4]. Time permitting, I may try to improve the accuracy of the algorithm, especially for moving subjects.

Technical Approach

Detecting heart rate in video consists of three main steps. First, video clips of various faces must be collected. Second, the facial region must be detected and tracked across each frame of the video since the face is the only portion of the frame that will contain heart rate information. Third, the plethysmographic signal must be extracted from the change in pixel colors within the face region over time and analyzed to determine the prominent frequency within the heart rate range.

Data Collection

I will collect video data of a few subjects with varying skin-tones and in varying lighting conditions. Each video will be approximately one minute in length and will be taken with a MacBook Pro webcam. Multiple videos of each subject will be taken, some with the subject still, some with the subject moving slightly, and possibly some after the subject has exercised to check heart rate detection at a wider range.

During video recording, subjects will be wearing a commercial heart rate monitor so that the heart rate results from the video can be compared to a ground truth.

Face Detection and Tracking

Face detection and tracking will be performed using Haar cascades classifiers as proposed by Viola and Jones [6] and improved by Lienhart *et*

al. [3]. The classifiers consist of several simpler classifiers applied in stages to a region of interest (ROI). Each simple classifier contains weighted votes of multiple basic classifiers trained on Haar-like features designed to detect edges, lines, and blobs. I will use the OpenCV Cascade Classifier for this face detection [2]. The cascade classifier will be pre-trained on positive and negative frontal face images. To detect the face in my video frames, I will slide the classification window across the image frames at varying sizes.

To maintain consistency across frames, if no face is detected in a frame, I will use the face from the previous frame, and if multiple faces are detected, I will use the face nearest to that in the previous frame. The face detector will output a bounding box for the face. Initially, I will use the center 60% of the bounding box width and the full height as my ROI, as was done by Poh *et al.* [4].

Time-permitting, I may try to improve the face detection and tracking and the estimation of the ROI. I could experiment with detecting other features on the face (eyes for example) and using this as the determination for which face to choose if there are multiple faces detected. I could improve the chosen ROI by segmenting out the facial region within the bounding box to avoid including hair or background in my ROI. I could also remove a region containing the eyes to reduce the non-skin pixels within the ROI, or experiment with segmenting out only the forehead region since that was shown to have strongest plethysmographic signal. Finally, I could try to detect large head movements and ignore the faces found in these frames to reduce the error due to motion.

Heart Rate Detection

Once I have an ROI for each frame, I can begin to extract the heart rate from the color image data. The first step will be to average the pixels in the ROI across each color channel to get three signals $x_R(t)$, $x_G(t)$, and $x_B(t)$ corresponding to the average red, green, and blue facial pixels at time

t . I will then normalize these signals across a 30-second sliding window with a 1-second stride (so the heart rate will be re-estimated every second).

I will then use ICA to extract the independent source signals from the observed mixed color signals. ICA assumes that the number of source signals is no more than the number of observed signals, so I will assume three source signals $s_1(t)$, $s_2(t)$, and $s_3(t)$ contributing to the observed color changes in the three channels. ICA assumes the observed mixed signals are a linear combination of these source signals. Although this assumption may not be valid as changes in blood volume and the intensity of reflected light in skin tissue over distance may be nonlinear, for the 30-second time window it should be a reasonable approximation. With this linear approximate for signal mixing, we have

$$x(t) = As(t)$$

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

where $x(t) = [x_R(t) \ x_G(t) \ x_B(t)]^T$, $s(t) = [s_1(t) \ s_2(t) \ s_3(t)]^T$, and A is a 3x3 matrix of coefficients. Then ICA attempts to find an approximation of A^{-1} that maximizes the non-Gaussianity of each source. I will use FastICA to recover the approximate source signals $s(t)$.

Poh *et al.* found that the second source signal usually had the strongest plethysmographic signal although there is no ordering to the source signals from ICA [4]. I will start by using this second source signal, but I may experiment with other methods such as choosing the source signal with the highest peak or detecting the heart rate from each of the signals and choosing the one that is most consistent with previous measurements.

Once I have the somewhat isolated plethysmographic signal, I can use Fourier transform to examine its frequency spectrum and determine the prominent signal frequencies. I can isolate frequency peaks in the power spectrum within the range 0.75 to 4 Hz, which corresponds to physiological heart rate ranges of 45 to 240 bpm. Following Poh *et al.*, I will only select spikes within

0.2 Hz of the previous measurement as heart-rates generally change less than 12 bpm over a single second [4].

Milestones

Borrow or order a HR monitor	May 4
Face tracking	May 9
Progress report	May 11
Color signal processing and ICA	May 16
Fourier transform and heart rate extraction	May 23
Potential algorithm fine-tuning and improvements	May 27
Project presentation	June 1
Project paper	June 6

References

- [1] G. Balakrishnan, F. Durand, and J. Guttag. Detecting pulse from head motions in video. *CVPR*, 2013.
- [2] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [3] R. Leinhardt, A. Kuranov, and V. Pisarevsky. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. *MRL Technical Report*, 2002.
- [4] M.-Z. Poh, D. McDuff, and R. Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optical Society of America*, 2010.
- [5] W. Verkrusse, L. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Opt Express*, 2008.
- [6] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *Accepted Conference on Computer Vision and Pattern Recognition*, 2001.
- [7] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. *SIGGRAPH*, 2012.