Heart Rate Spy

Nolan Hergert
Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, Pennsylvania
Email: nhergert@andrew.cmu.edu

William Keyes
Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, Pennsylvania
Email: wkeyes@andrew.cmu.edu

***Abstract*—Measuring heart rate traditionally requires special equipment and physical contact with the subject. Previous work has shown that consumer-grade web cameras can provide useful signals for remote heart rate measurement. We present our implementation of a system that allows automated measurement of heart rate using a standard web camera and laptop computer. We also discuss environmental factors which presented difficulties for our methods and compare our methods to previously published results.**

I. Introduction

Heart rate (HR) is an important indicator of the general health of a patient. Further, it is a useful measure for doctors to have available when determining a proper course of treatment. One standard method for determining heart rate, particularly in patients with cardiac irregularities, is the electrocardiogram (ECG). While very accurate, ECGs requires specialized equipment operated by trained technicians and as such are not practical outside of dedicated care centers.

Developments in photoplethysmography (PPG) have helped reduce dependence on such specialized equipment for routine measurement of heart rate. PPG refers to a set of techniques for measuring vital signals and organ function using non-invasive light emitters and receivers. Like ECG, PPG traditionally requires dedicated hardware, here placed on a subject's finger or earlobe. While non-invasive and less complicated than ECG, these forms of PPG still require physical contact with the subject and the necessary hardware may not be easily accessible.

Recently, other research groups have demonstrated that video recorded with a consumer-grade web camera can be used as an input signal for PPG. This is an important development for several reasons. First, standard PPG devices often use specialized light sources and detectors in the infra-red (IR) or near-infra-red range. Second, web cameras are now commodity hardware, particularly on modern laptop computers; an accurate and robust software PPG solution for measuring heart rate would have a large potential user base.

We present a system that builds on and combines ideas developed by Verkruysse, et al. [1] and Poh, et al. [2]. We believe our system finds a good balance between the simplicity of the first approach and the robustness of the second. While the outcome is not at a level suitable for use by the general public, it is a good demonstration of the basic principles and a reaffirmation that simple techniques can lead to surprisingly strong results.

A. Principles

Because skin is translucent, many properties of the light reflected off the body are determined by subcutaneous factors. PPG, at least for heart rate measurement, is concerned with the volume of blood in the veins in a region of skin at any given time. As the blood volume pulse travels through the body, the reflectance and transmittance of light oscillates with the same period as the heart rate [3].

Finger-tip sensors typically measure the blood volume pulse by measuring how light in the IR or near-IR spectrum is transmitted through the finger as a function of time. It is also possible to measure changes in reflected light [3]. This is the approach favored by our system and other systems based on unmodified consumer cameras. In particular, we examine variations in the red, green, and blue (RGB) color channels returned by the camera.

Previous work by both Verkruysse, et al. [1] and Poh, et al. [2] is fundamentally based on spatial averaging of pixel values in a region of interest (ROI) in a video. However the methods of selecting the ROI and the additional processing that is performed after averaging differ.

B. Fixed ROI and Frequency Analysis

Verkruyssee, et al. recorded video of subjects lying down or sitting upright, then manually selected a ROI from a still frame of the video. After the pixel values in the ROI were spatially averaged, the discrete Fourier transform (DFT) of the variance signal was used to determine the frequency with highest power. This frequency was chosen as the heart rate. Applying this technique to each color channel individually determined that the green channel variance most often contained the strongest heart rate signal.

C. Dynamic ROI and Independent Component Analysis

Poh, et al. took recorded video or live streams of subject engaged in normal computer use and automatically selected a ROI using the face detection routines provided by the OpenCV [4] computer vision library. The spatial average pixel values for all three color channels were computed and independent component analysis (ICA) was applied to the variance signals. ICA is a method of blind source separation (BSS), in which the measured signals are assumed to be the result of mixing some set of source signals using an unknown system. BSS attempts to reconstruct the original source signals using the measured signals and a set of constraints. One of the resultant signals from ICA was selected as containing the strongest heart rate signal. The heart rate was chosen using a frequency-domain method similar to the one described above.

II. Methods

Our application runs in real-time on a 2009 Apple MacBook Pro laptop and was built using Java, Processing [5], and OpenCV. We use the integrated iSight camera, which provides 24-bit RGB data and allows us to processes 640 pixel by 480 pixel video at 30 frames per second. Most testing was performed using streaming video, although we recorded some video segments for more detailed off-line analysis.

Given the k^{th} input frame $F[k]$, we first detect faces using the Viola-Jones object detection framework as implemented in OpenCV [6]. This provides a list of bounding boxes which surround faces in the image, from which we select the largest. If this is the first time a face appears or if a new ROI is requested, we define the ROI R to be a rectangle $5/9$ the width and $3/4$ the height of the bounding box, centered in the bounding box. This focuses our efforts on the area of the face with the most exposed skin and helps avoid errors caused by involuntary subject movement and boundary conditions such as hair or the edges of the face. R has constant dimensions, but its position in the image changes based on updated face coordinates.

New face coordinates are integrated into the existing coordinates using an exponentially-weighted moving average (EMA) filter

$$y[n] = \tau x[n] + (1 - \tau)y[n - 1] \tag{1}$$

where $y[n]$ is the averaged value at time n, $x[n]$ is the input at time n, and $0 < \tau < 1$ is the time-constant of the filter. We use $\tau = 1/100$ for updating the location of R.

The frame $F[k]$ is composed of three color channels; let $F_R[k]$, $F_G[k]$, and $F_B[k]$ represent the red, green, and blue channels respectively. Define the instantaneous average pixel value of the green channel, $g[k]$ to be

$$g[k] = \frac{\sum_{i,j \in R} F_G[k](i, j)}{\text{Area}(R)} \tag{2}$$

where $F_G[k](i, j)$ is the value of the pixel in $F_G[k]$ at location (i, j). The values $r[k]$ and $b[k]$ are defined analogously for the red and blue channels, but we are primarily concerned with the green channel.

Define the instantaneous variance of the green channel, $v_G[k]$ to be

$$v_G[k] = g[k] - \bar{g}[k] \tag{3}$$

where $\bar{g}[k]$, the time-average pixel value, is the result of applying an EMA filter ($\tau = 1/50$) to the instantaneous pixel values. To reduce variance noise resulting from sensor noise in the camera, we filtered $v_G[k]$ with an EMA filter ($\tau = 1/3$) to produce $\bar{v}_G[k]$.

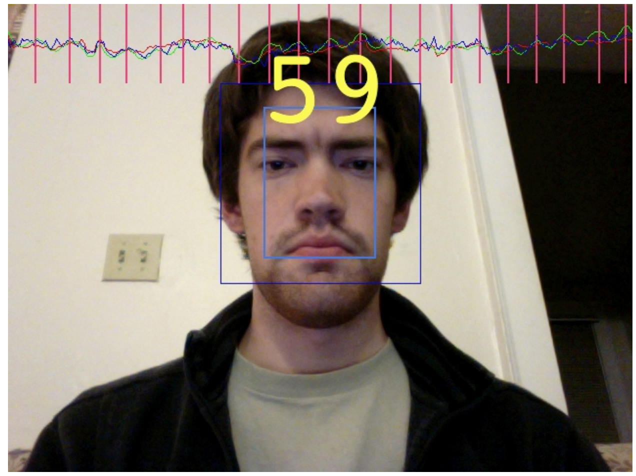


Fig. 1. The application interface

Heart rate was determined by measuring peak-to-peak times in $\bar{v}_G[k]$. A value is marked as a peak if it preceded by and followed by a "valley" that is at least a fixed amount lower than the possible peak value [7]. Peaks that would cause abnormal or sudden changes in HR were discarded, and the last four peak-to-peak timings were averaged to get the final heart rate.

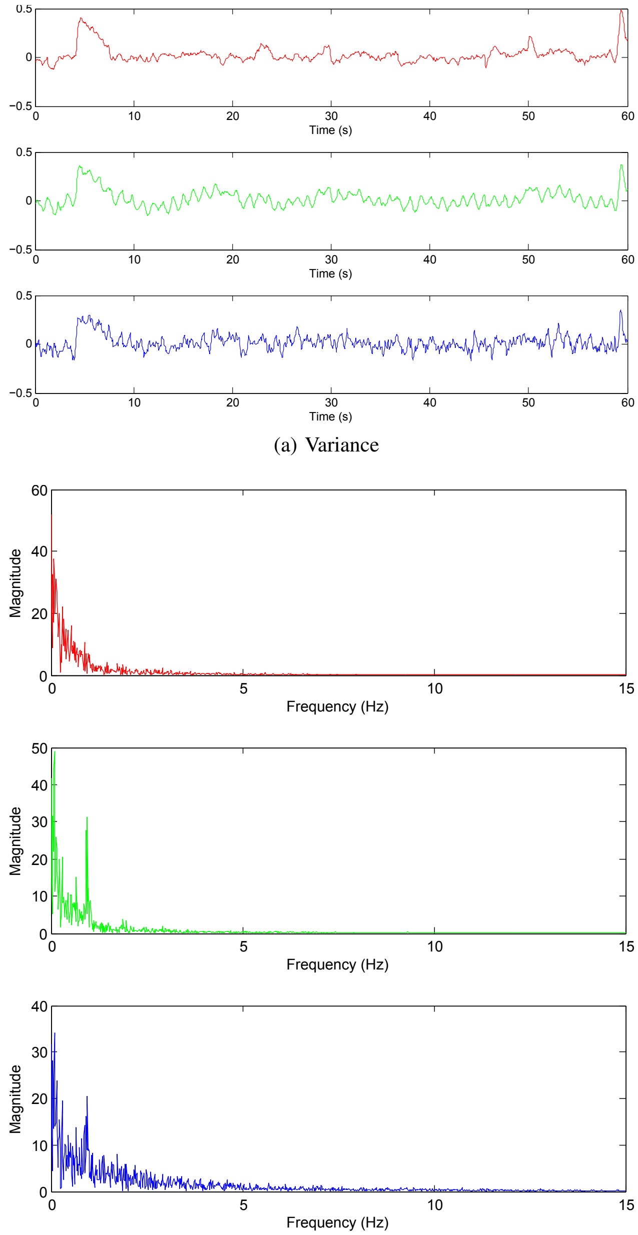
The application consists of a single window which displays $\bar{v}_G[k]$ annotated with detected peaks, the calculated heart rate, and the ROI used for computation. Users can press a key to calculate a new ROI; this is required after sudden, large movements or when a new user first appears in front of the camera.

III. Results

An example of computed variance signals and their corresponding frequency plots are shown in Figure 2. We see that the red channel variance is mostly noise, while the green channel has clear periodic oscillations. The blue channel also has oscillations, but they are not as clean as those in the the green channel. Examining the frequency plots confirms our observations of the time signal; the green channel has a clear peak slightly below 1 Hz, which is typical for a resting heart rate.

Peak sizes in the variance signal changed with lighting conditions and skin tone variation. We found that our system had little difficulty adjusting to different skin tones but that poor lighting could significantly degrade results. Natural or incandescent light provided the best results, most likely because both sources have components in the IR and near IR spectrum. While our system still functioned under primarily fluorescent lighting, it was more susceptible to noise and false positives and negatives in the peak detector.

To test the accuracy of our system, we compared our heart rate measurements to those determined by a commercially available fingertip pulse oximeter. When stable, our system consistently produced heart rate measurements within a 5 to 6 beat per minute window around the reading from the commercial sensor. However, our measured value fluctuated



(a) Variance

(b) Variance Spectra

Fig. 2. Variance plots for the three color channels (a) and the FFT of each variance signal (b). Note the large peak in the green channel and the corresponding, but smaller, peak in the blue channel.

more often and more rapidly than the reference measurement. We suspect this is mostly caused by the sensitivity of our peak detection scheme to false-positive peaks and could be avoided with more aggressive filtering or linear prediction.

We also attribute some of this variance to changes in frame rate. We assume video captured at a constant 30 frames per second when computing the heart rate, but the actual frame rate will vary up and down depending on the demands placed on the computer. Reducing such variations requires a careful balance, as the ability to respond quickly to real heart rate changes is a desirable feature of a measurement system.

Our system is also very sensitive to subject movement. While our method for determining a ROI is fairly resistant to involuntary motion, such as eye blinks, breathing, and small movements of the head, we cannot successfully respond to larger movements such as looking up or down, or moving from side to side. This was expected given the measures that Verkruyssee, et al. took to eliminate motion when using spatial averaging as the primary analysis technique. We believe that motion tolerance can be increased by tracking the face more accurately, perhaps by using a feature tracking algorithm in addition to standard face detection. Though there is likely a limit to the motion tolerance that can be achieved using our averaging technique without resorting to techniques like those used by Poh, et al., we believe that the simplicity of our technique has advantages for use in embedded or other low cost devices.

IV. Future Work and Applications

As we have demonstrated, implementing a system which uses a standard web camera to measure heart rate is both possible and practical on commodity hardware. However, as a proof-of-concept system, there are many areas in which it could be improved or expanded.

A. Improvements

As discussed above, there are many factors that can cause the system to produce bad readings. Any initial improvements should be directed at increasing robustness. We would like to further experiment with methods for reducing face bounding box variability, possibly by tracking arbitrary feature points or by using a correlation filter within the face region.

While our method of peak detection works well and avoids the need to the fast Fourier transform, our current filtering methods allow misclassified peaks to greatly distort the resultant heart beat. Though our current approach of discarding bad peaks and averaging could be extended and refined, we believe a more advanced linearly predictive filtering technique would give much better results. This would allow us to better maintain a heart rate measurement in the absence of peaks or presence of extra peaks, while still tracking the actual heart rate when presented with enough consistent peaks.

If more advanced processing techniques are allowed, we suspect that ICA, as demonstrated by Poh, et al., or principle component analysis (PCA) will provide superior illumination and motion invariance. PCA is particularly promising, as it is designed to find the representation of a data set that maximizes variance, our measure of heart rate. In our initial trials with PCA, it was unclear which combination of the resulting basis vectors best represented the variance that results from the blood volume pulse. More experimentation is required to discover how PCA can be applied most effectively.

B. Applications

The simple techniques required for our system make it particularly promising for embedded devices. In fact, Phillips Electronics recently released an application for the Apple iPad 2 tablet computer that appears to employ similar techniques to measure heart rate [8]. In addition to the potential for simple mobile or remote heart rate measurement, we also believe these techniques could be useful as a liveness measure in biometric access control systems such as face recognizers.

Other research has shown that combining an IR light source with a modified camera can allow the visualization of blood vessels and other subcutaneous structures [9]. Combining this technique with our system for heart rate measurement would produce an integrated and powerful diagnosis tool.

V. Conclusion

We have shown that it is possible and practical to use a consumer-grade web camera and commodity hardware to create a system that can measure the heart rate of a subject remotely and using available light. Although our technique is susceptible to interference from motion and illumination change, it successfully automates Verkruysse, et al.'s spatial averaging method and illustrates that the more powerful methods of Poh, et al. are not necessary to produce useful results. Although there is much work that can be done to improve the system, it and other photoplethysmographic techniques based on consumer hardware show promise for cost-efficiently increasing the availability and quality of medical care.

Acknowledgment

The authors would like to thank Professor Marios Savvides, Utsav Prabhu, and Andres Rodriquez for their support in this project and in the rest of 18-794, *Pattern Recognition*.

References

[1] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Express*, vol. 16, no. 26, pp. 21 434–21 445, Dec 2008. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-16-26-21434

[2] 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, pp. 10 762–10 774, May 2010. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-10-10762

[3] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological Measurement*, vol. 28, no. 3, p. R1, 2007. [Online]. Available: http://stacks.iop.org/0967-3334/28/i=3/a=R01

[4] http://opencv.willowgarage.com/.

[5] http://processing.org.

[6] "Haar feature-based cascade classifier for object detection," http://opencv.itseez.com/modules/objdetect/doc/cascade_classification.html.

[7] E. Billauer, "peakdet: Peak detection using MATLAB," http://billauer.co.il/peakdet.html.

[8] Phillips, "Philips vital signs camera," http://www.vitalsignscamera.com/.

[9] V. P. Zharov, S. Ferguson, J. F. Eidt, P. C. Howard, L. M. Fink, and M. Waner, "Infrared imaging of subcutaneous veins," *Lasers in Surgery and Medicine*, vol. 34, no. 1, pp. 56–61, 2004. [Online]. Available: http://dx.doi.org/10.1002/lsm.10248