Detecting Pulse from Head Motions in Video

Guha Balakrishnan, Fredo Durand, John Guttag
MIT CSAIL
{balakg, fredo, guttag}@mit.edu

Abstract

We extract heart rate and beat lengths from videos by measuring subtle head motion caused by the Newtonian reaction to the influx of blood at each beat. Our method tracks features on the head and performs principal component analysis (PCA) to decompose their trajectories into a set of component motions. It then chooses the component that best corresponds to heartbeats based on its temporal frequency spectrum. Finally, we analyze the motion projected to this component and identify peaks of the trajectories, which correspond to heartbeats. When evaluated on 18 subjects, our approach reported heart rates nearly identical to an electrocardiogram device. Additionally we were able to capture clinically relevant information about heart rate variability.

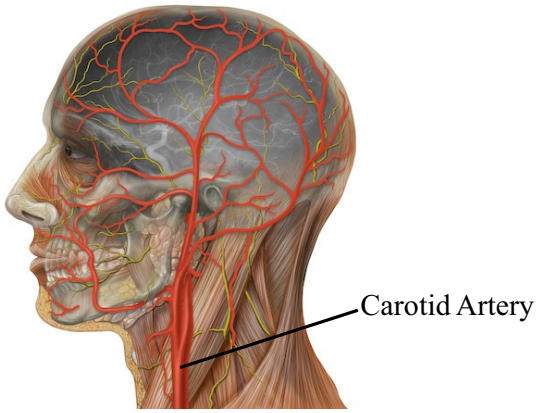

Figure 1: Blood flows from the heart to the head via the carotid arteries on either side of the head [11].

1. Introduction

Heart rate is a critical vital sign for medical diagnosis. There is growing interest in extracting it without contact, particularly for populations such as premature neonates and the elderly for whom the skin is fragile and damageable by traditional sensors. Furthermore, as the population ages, continuous or at least frequent monitoring outside of clinical environments can provide doctors with not just timely samples but also long-term trends and statistical analyses. Acceptance of such monitoring depends in part on the monitors being non-invasive and non-obtrusive.

In this paper, we exploit subtle head oscillations that accompany the cardiac cycle to extract information about cardiac activity from videos. In addition to providing an unobtrusive way of measuring heart rate, the method can be used to extract other clinically useful information about cardiac activity, such as the subtle changes in the length of heartbeats that are associated with the health of the autonomic nervous system.

The cyclical movement of blood from the heart to the head via the abdominal aorta and the carotid arteries (Fig. 1) causes the head to move in a periodic motion. Our algorithm detects pulse from this movement. Our basic approach is to track feature points on a person's head, filter their velocities by a temporal frequency band of interest, and use principal component analysis (PCA) to find a periodic signal caused by pulse. We extract an average pulse rate from this signal by examining its frequency spectrum and obtain precise beat locations with a simple peak detection algorithm.

Our method is complementary to the extraction of pulse rate from video via analysis of the subtle color changes in the skin caused by blood circulation [14, 18]. These methods average pixel values for all channels in the facial region and temporally filter the signals to an appropriate band. They then either use these signals directly for analysis [18] or perform ICA to extract a single pulse wave [14]. They find the frequency of maximal power in the frequency spectrum to provide a pulse estimate. Philips also produced a commercial app that detects pulse from color changes in real-time [13]. These color-based detection schemes require that facial skin be exposed to the camera. In contrast our approach is not restricted to a particular view of the head, and is effective even when skin is not visible. There have also been studies on non-invasive pulse estimation using modalities other than video such as thermal imagery [6] and photoplethysmography (measurement of the variations in trans-

mitted or reflected light in the skin) [21].

The analysis of body motion in videos has been used in different medical contexts, such as the measurement of respiration rate from chest movement [17, 13], or the monitoring of sleep apnea by recognizing abnormal respiration patterns [20]. Motion studies for diseases include identification of gait patterns of patients with Parkinson's disease[4], detection of seizures for patients with epilepsy [12] and early prediction of cerebral palsy [1]. The movements involved in these approaches tend to be larger in amplitude than the involuntary head movements due to the pulse.

Our work is also inspired by the amplification of imperceptible motions in video [22, 10]. But whereas these methods make small motions visible, we want to extract quantitative information about heartbeats.

The idea of exploiting Newton's Third Law to measure cardiac activity dates back to at least the 1930's, when the ballistocardiogram (BCG) was invented [15]. The subject was placed on a low-friction platform, and the displacement of the platform due to cardiac activity was measured. The BCG is not widely used anymore in clinical settings. Other clinical methods using a pneumatic chair and strain-sensing foot scale have also been successful under laboratory conditions[9, 8]. Ballistocardiographic head movement of the sort studied here has generally gained less attention. Such movement has been reported during studies of vestibular activity and as an unwanted artifact during MRI studies [2]. Recently, He *et al.*[7] proposed exploiting head motion measured by accelerometers for heart rate monitoring as a proxy for traditional BCG.

In this paper we first describe a novel technique that extracts a pulse rate and series of beat sequences from video recordings of head motion. We then evaluate our system's heart rate and beat location measurements on subjects against an electrocardiogram. Results show that our method extracts accurate heart rates and can capture clinically relevant variability in cardiac activity.

2. Background

2.1. Head Movement

The head movements related to cardiac activity are small and mixed in with a variety of other involuntary head movements. From a biomechanical standpoint, the head-neck system and the trunk can be considered as a sequence of stacked inverted pendulums. This structure allows the head unconstrained movement in most axes. There are several sources of involuntary head movement that complicate the isolation of movements attributable to pulsatile activity. One is the pendular oscillatory motion that keeps the head in dynamic equilibrium. Like He *et al.* [7], we found that the vertical direction is the best axis to measure the movement of the upright head caused by pulse because the horizontal axis tends to capture most of the dynamic equilibrium swaying. A second source of involuntary head movement is the bobbing caused by respiration. We deal with this by filtering out low-frequency movement.

The net acceleration of involuntary vertical head movement has been measured to be around 10 mG ($\approx .098\frac{m}{s^2}$) [7]. The typical duration of the left ventricular ejection time of a heart cycle is approximately $\frac{1}{3}$ seconds. Using these numbers we can calculate a rough estimate of the displacement of head movement to be $\frac{1}{2} \cdot 0.098 \cdot (\frac{1}{3})^2 \approx 5$ mm. Though this calculation neglects the complex structure of the head system, it does provide an indication of how small the movement is.

2.2. Beat-to-beat Variability

Pulse rate captures the average heart rate over a period of time (e.g., 30 seconds). It is useful primarily for detecting acute problems. There is a growing body of evidence [16] that measuring beat-to-beat variations provides additional information with long-term prognostic value. The most established of these measures is heart rate variability (HRV). HRV measures the variation in the length of individual normal (sinus) heartbeats. It provides an indication of the degree to which the sympathetic and parasymathetic nervous systems are modulating cardiac activity. To measure HRV, the interarrival times of beats must be accurately measured, which can be determined by locating the "R" peaks in successive beats in an electrocardiogram (ECG). A lack of sufficient variation when the subject is at rest suggests that the nervous system may not perform well under stress. Patients with decreased HRV are at an increased risk of adverse outcomes such as fatal arrhythmias.

3. Method

Our method takes an input video of a person's head and returns a pulse rate as well as a series of beat locations which can be used for the analysis of beat-to-beat variability. We first extract the motion of the head using feature tracking. We then isolate the motion corresponding to the pulse and project it onto a 1D signal that allows us to extract individual beat boundaries from the peaks of the trajectory. For this, we use PCA and select the component whose temporal power spectrum best matches a pulse. We project the trajectories of feature points onto this component and extract the beat locations as local extrema.

Fig. 2 presents an overview of the technique. We assume the recorded subject is stationary and sitting upright for the duration of the video. We start by locating the head region and modeling head motion using trajectories of tracked feature points. We use the vertical component of the trajectories for analysis. The trajectories have extraneous motions at frequencies outside the range of possible pulse rates, and so we temporally filter them. We then use

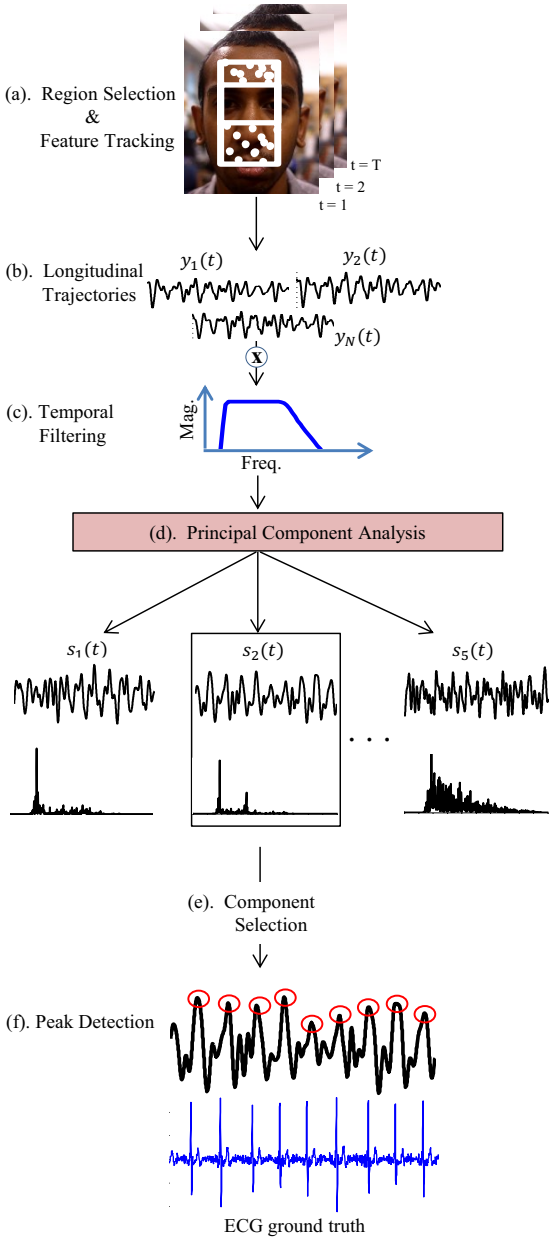

Figure 2: Overview of our pulse estimation approach. (a) A region is selected within the head and feature points are tracked for all frames of the video. (b) The vertical component is extracted from each feature point trajectory. (c) Each trajectory is then temporally filtered to remove extraneous frequencies. (d) PCA decomposes the trajectories into a set of source signals s_1, s_2, s_3, s_4, s_5. (e) The component which has clearest main frequency is selected. (f). Peak detection identifies the beats of the signal.

PCA to decompose the trajectories into a set of independent source signals that describe the main elements of the head motion. To choose the correct source for analysis and computation of the duration of individual beats, we examine the frequency spectra and select the source with the clearest main frequency. Average pulse rate is identified using this frequency. For more fine-grained analysis and calculation of beat durations, we perform peak detection in the time-domain.

3.1. Region Selection and Tracking

We find a region of interest containing the head and track feature points within the region. For videos where the front of the face is visible, we use the Viola Jones face detector [19] from OpenCV 2.4 [3] to first find a rectangle containing the face. We opt to use the middle 50% of the rectangle widthwise and 90% heightwise from top in order to ensure the entire rectangle is within the facial region. We also remove the eyes from the region so that blinking artifacts do not affect our results. To do this we found that removing the subrectangle spanning 20% to 55% heightwise works well. For videos where the face is not visible, we mark the region manually.

We measure the movement of the head throughout the video by selecting and tracking feature points within the region. We apply the OpenCV Lucas Kanade tracker between frame 1 and each frame $t = 2 \cdots T$ to obtain the location time-series $\langle x_n(t), y_n(t) \rangle$ for each point n. Only the vertical component $y_n(t)$ is used in our analysis. Since a modern ECG device operates around 250 Hz to capture heart rate variability and our videos were only shot at 30 Hz, we apply a cubic spline interpolation to increase the sampling rate of each $y_n(t)$ to 250 Hz.

Many of the feature points can be unstable and have erratic trajectories. To retain the most stable features we find the maximum distance (rounded to the nearest pixel) traveled by each point between consecutive frames and discard points with a distance exceeding the mode of the distribution.

3.2. Temporal Filtering

Not all frequencies of the trajectories are required or useful for pulse detection. A normal adult's resting pulse rate falls within $[0.75, 2]$ Hz, or $[45, 120]$ beats/min. We found that frequencies lower than 0.75 Hz negatively affect our system's performance. This is because low-frequency movements like respiration and changes in posture have high amplitude and dominate the trajectories of the feature points. However, harmonics and other frequencies higher than 2 Hz provide useful precision needed for peak detection. Taking these elements into consideration, we filter each $y_n(t)$ to a passband of $[0.75, 5]$ Hz. We use a 5^{th} order butterworth filter for its maximally flat passband.

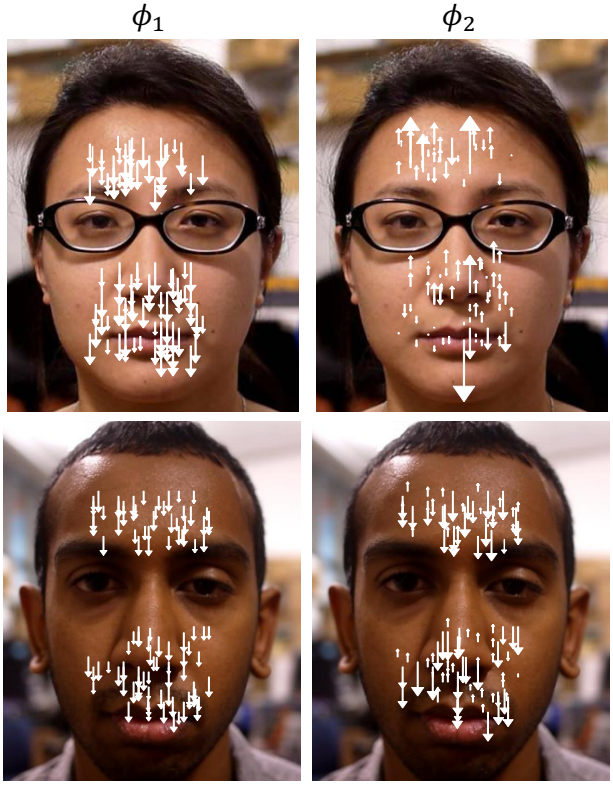

Figure 3: Examples of the first two eigenvectors for two subjects. Each white arrow on a face represents the magnitude and direction of a feature point's contribution to that eigenvector. The eigenvector decomposition is unique to each subject.

3.3. PCA Decomposition

The underlying source signal of interest is the movement of the head caused by the cardiovascular pulse. The feature point trajectories are a mixture of this movement as well as other motions caused by sources like respiration, vestibular activity and changes in facial expression. We must decompose this mixed motion into subsignals to isolate pulse. To do this we consider the multidimensional position of the head at each frame as a separate data point and use PCA to find a set of main dimensions along which the position varies. We then select a dimension on which to project the position time-series to obtain the pulse signal.

Formally, given N feature points, we represent the N-dimensional position of the head at frame t as $m_t = [y_1(t), y_2(t), \cdots, y_N(t)]$. The mean and the covariance matrix of the positions are:

$$\bar{m} = \frac{1}{|T|}\sum_{i=1}^{T} m_i \tag{1}$$

$$\Sigma_m = \frac{1}{T}\sum_{i=1}^{T}(m_t - \bar{m})(m_t - \bar{m})^T \tag{2}$$

PCA finds the principal axes of variation of the position as the eigenvectors of the covariance matrix:

$$\Sigma_m \Phi_m = \Phi_m \Lambda_m \tag{3}$$

where Λ_m denotes a diagonal matrix of the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_N$ corresponding to the eigenvectors in the columns of Φ_m, $\phi_1, \phi_2, \cdots, \phi_N$. Fig. 3 displays the first two eigenvectors for two of the subjects. Each eigenvector represents the N-dimensional direction and magnitude of movement for the feature points. The eigenvectors differ for each subject. We obtain the 1-D position signal $s_i(t)$ by projecting the position time-series onto ϕ_i:

$$s_i(t) = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_T \end{pmatrix} \cdot \phi_i \tag{4}$$

There are periods in the video during which the head moves abnormally (e.g. swallowing, adjustments in posture). Such movement adds variance to the position vectors, thereby affecting the PCA decomposition. To deal with this one could discard a percentage α of the m_t with the largest $L2$-norms before performing PCA. However, all of the m_t must still be used in the projection step (Eq. 4) to produce a complete signal. We set α at a value of 25% for our experiments.

A popular alternative to PCA is independent component analysis (ICA). We did not see any improvement in our results using ICA.

3.4. Signal Selection

The question remains of which eigenvector to use for pulse signal extraction. The eigenvectors are ordered such that ϕ_1 explains the most variance in the data, ϕ_2 explains the second most, and so on. Although ϕ_1 explains most of the variance, s_1 may not be the clearest pulse signal for analysis. We instead choose the s_i that is most periodic. We quantify a signal's periodicity as the percentage of total spectral power accounted for by the frequency with maximal power and its first harmonic.

We found that it was not necessary to consider any signals beyond the first five, i.e. $s_1, ..., s_5$ for any of our subjects. We label the maximal frequency of the chosen signal f_{pulse} and approximate the pulse rate as $\frac{60}{f_{pulse}}$ beats per minute.

3.5. Peak Detection

Average pulse rate alone is not sufficient to fully evaluate the cardiovascular system. Clinicians often assess beat-to-beat variations to form a complete picture. To allow for

such analysis, we perform peak detection on the selected PCA component signal.

The peaks are close to $\frac{1}{f_{pulse}}$ seconds apart with some variability due to the natural variability of heartbeats, variations of the head motion, and noise. We label each sample in the signal as a peak if it is the largest value in a window centered at the sample. We set the length of the window (in samples) to be $round(\frac{f_{sample}}{f_{pulse}})$, where $f_{sample} = 250$Hz.

4. Experiments

We implemented our approach in MATLAB. Videos were shot with a Panasonic Lumix GF2 camera in natural, unisolated environments with varying lighting. All videos had a frame rate of 30 frames per second, 1280 x 720 pixel resolution and a duration of 70-90 seconds. We connected subjects to a wearable ECG monitor [5] for ground truth comparison. This device has a sampling rate of 250 Hz and three electrodes that we placed on the forearms.

4.1. Visible Face

We extracted pulse signals from 18 subjects with a frontal view of the face (as in Fig. 3). The subjects varied in gender (7 female, 11 male) and skin color. They ranged from 23-32 years of age and were all seemingly healthy. We calculate our program's average pulse rate using the frequency of maximal power for the selected PCA component. Similarly, we compute the true pulse rate by finding the main frequency of the ECG spectrum. Table 1 presents our results. The average rates are nearly identical to the true rates for all subjects, with a mean error of 1.5%. The number of peaks were also close to ground truth values, with a mean error of 3.4%.

We also evaluate the ability of our signal to capture subtle heart rate variability. Clinically meaningful HRV measures typically use 10-24 hours of ECG data. Therefore we did not attempt to compute any of these for our 90 second videos. Instead, we compare the distributions of time between successive peaks for each signal. Incorrect or missed peaks can introduce spurious intervals too large or small to be caused by the natural variations of the heart. We account for these cases by only considering intervals with length within 25% of the average detected pulse period.

We use the Kolmogorov-Smirnov (KS) test to measure the similarity of the distributions, with the null hypothesis being that the observations are from the same distribution. Table 2 presents the results. At a 5% significance level, 16 of the 18 pairs of distributions were found to be similar. Fig. 4 presents histograms of 4 of the 16 distributions binned at every 0.05 seconds. Our method was able to capture a wide range of beat-length distributions shapes, from the flat distribution of subject 4 to the peakier distribution of subject 10.

Table 1: Average pulse rate and number of peaks detected from ECG and by our method.

	Avg. Pulse (beats per minute)		Number of peaks	
Sub.	ECG	Motion (% error)	ECG	Motion(% error)
1	66.0	66.0(0)	99	98(1.0)
2	54.7	55.3(1.1)	82	84(2.4)
3	81.3	82.6(1.6)	122	116(4.9)
4	44.7	46.0(2.9)	67	70(4.5)
5	95.3	96.0(0.7)	143	142(0.7)
6	78.9	78.0(1.1)	92	78(15.2)
7	73.3	71.3(2.7)	110	100(9.1)
8	59.3	58.6(1.2)	89	88(1.1)
9	56.7	58.6(3.4)	85	84(1.2)
10	78.7	79.3(0.8)	118	117(0.8)
11	84.7	86.6(2.2)	127	121(4.7)
12	63.3	62.6(1.1)	95	95(0)
13	59.3	60.0(1.2)	89	89(0)
14	60.0	61.3(2.2)	90	89(1.1)
15	80.0	81.3(1.6)	120	114(5.0)
16	74.7	74.6(0.1)	112	110(1.8)
17	50.0	49.3(1.4)	75	76(1.3)
18	77.1	78.8(2.2)	90	85(5.6)

Table 2: Results when comparing the interpeak distributions of the ECG and our method. Presented are the means (μ) and standard deviations (σ) of each distribution, the number of outliers removed from our distribution, and the p-value of distribution similarity. 16 of the 18 pairs of distributions were not found to be significantly different.

Sub.	ECG $\mu(\sigma)$	Motion $\mu(\sigma)$	KS-Test p-value
1	0.91(0.06)	0.90(0.06)	0.89
2	1.08(0.08)	1.06(0.11)	0.52
3	0.73(0.04)	0.73(0.08)	0.05
4	1.34(0.19)	1.28(0.18)	0.14
5	0.62(0.03)	0.63(0.07)	<0.01
6	0.76(0.04)	0.76(0.04)	0.64
7	0.81(0.05)	0.81(0.06)	0.85
8	1.01(0.04)	1.02(0.09)	0.16
9	1.04(0.07)	1.04(0.11)	0.27
10	0.75(0.04)	0.75(0.04)	0.75
11	0.70(0.06)	0.70(0.08)	0.30
12	0.94(0.08)	0.94(0.09)	0.85
13	0.99(0.04)	0.98(0.12)	<0.01
14	0.99(0.11)	0.98(0.12)	0.47
15	0.74(0.05)	0.75(0.06)	0.95
16	0.80(0.05)	0.80(0.06)	0.60
17	1.18(0.08)	1.18(0.11)	0.70
18	0.76(0.05)	0.76(0.06)	0.24

4.1.1 Motion Amplitude

Pulse motion constitutes only a part of total involuntary head movement. We quantify the magnitude of the differ-

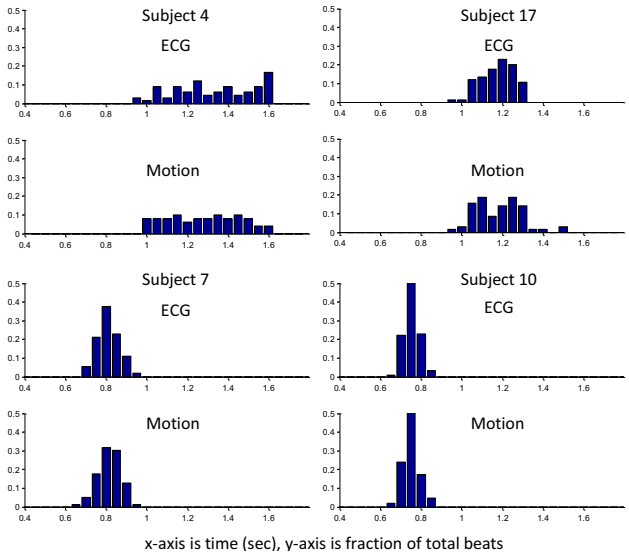

Figure 4: Beat distributions of the ECG and our motion method for 4 subjects. We were able to accurately capture a wide range of distribution shapes.

ent movements within $[0.75, 5]$ Hz by calculating root mean square (RMS) amplitudes of the feature point trajectories. For each subject we calculated the mean RMS amplitude of the trajectories before and after filtering to a passband within 5% of the pulse frequency. The mean RMS amplitude of the trajectories without filtering was 0.27 (std. dev of 0.07) pixels across the subjects. The mean RMS amplitude after filtering to the pulse frequency was 0.11 (0.05) pixels. Thus the pulse motion had roughly 40% the RMS amplitude of other head motions within the $[0.75, 5]$ Hz frequency range.

4.1.2 Comparison to Color-Based Detection and Noise Analysis

We compare the robustness of our method to a color-based pulse detection system [14] in the presence of noise. The color method spatially averages the R, G, and B channels in the facial area and uses independent component analysis (ICA) to decompose the signals into 3 independent source signals. The source with the largest peak in the power spectrum is then chosen as the pulse signal.

We added varying levels of zero-mean Gaussian noise to the videos and swept the standard deviation from 5 to 500 pixels. For each subject we found σ_{motion}, the maximum noise standard deviation before our method first produced an average pulse rate outside 5% of the true rate. We calculated σ_{color} in a similar manner for the color method. Fig. 5 plots the results. Our method outperformed color for 7 of the 17 subjects, and performed worse for 9 subjects. Note that color failed to produce a correct pulse rate for subject 7

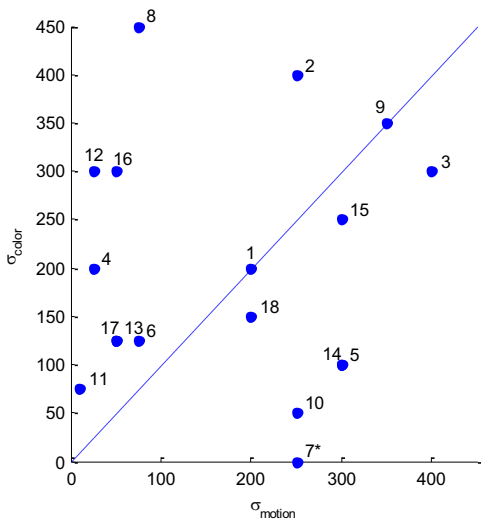

Figure 5: Comparison of our method to color-based detection. σ_{motion} and σ_{color} are the maximum noise standard deviations where motion and color give reliable results. The blue line is where $\sigma_{motion} = \sigma_{color}$. Our method worked longer for 7 of the 18 subjects while color worked longer for 9 subjects. The color method failed to give a correct result for subject 7 before the addition of noise.

before adding any noise.

We see a large variance in σ_{motion} and σ_{color} across the subjects, suggesting that there are subject-specific factors that affect performance. To understand why, we compare σ_{motion} against β, the ratio of the total energy of the feature points within 5% of f_{pulse} to the maximal energy at any other frequency. Fig. 6 plots σ_{motion} against β for all subjects. The subjects with the 10 highest σ_{motion} values also have 10 of the top 11 β values. This indicates that our method performs best for subjects with a large ballistocardiac motion relative to any other periodic head movement.

We were unable to find a similar relationship between σ_{color} and the frequency content of the R, G, and B channels. This is likely due to the layer of complexity introduced by the ICA algorithm. However, when simplifying the method to extracting a signal from the G channel alone, we found that noise performance is indeed strongly related to the ratio of power at the pulse frequency to the next largest power in the spectrum. Contrary to our initial hypothesis we saw no relationship between motion or color performance and skin tone.

4.2. Other Videos

One of the advantages of motion-based detection over color is that a direct view of the skin is not needed. We took videos of the backs of the heads of 11 subjects and a video of one subject wearing a mask, as shown in Fig. 7. We were

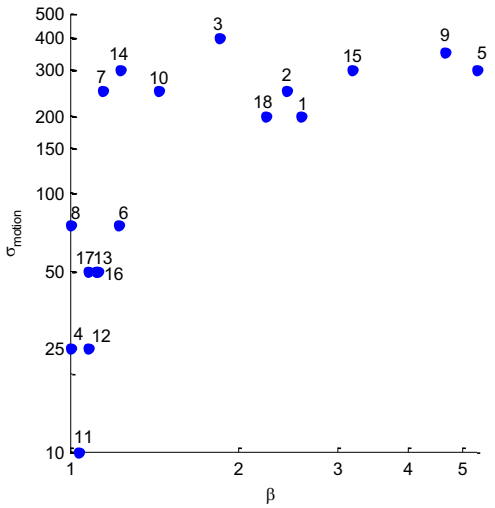

Figure 6: Log plot comparing σ_{motion} (the max noise standard deviation before our method produced an incorrect pulse rate) and β (ratio of the total energy of feature points at the pulse frequency to the maximal energy at any other frequency) for each subject. Subjects with large β tended to have better noise performance.

able to get average heart rates close to the true values for all videos.

We also tested our system on a 30-second video of a newborn recorded in situ at a hospital nursery (see Fig. 8). We also obtained a video of the baby's actual pulse rate from a hospital-grade monitor measuring the perfusion of blood to its skin. Our algorithm extracts a clean pulse signal that matches the pulse rate reported by the monitor.

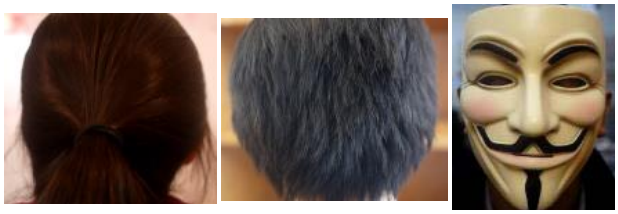

Figure 7: Reference frames from two videos of the back of the head and one of a face covered with a mask.

5. Discussion

Our results show it is possible to consistently obtain accurate pulse rate measurements from head motion. The results for beat detection were equally encouraging. Most of our beat interval distributions looked qualitatively similar

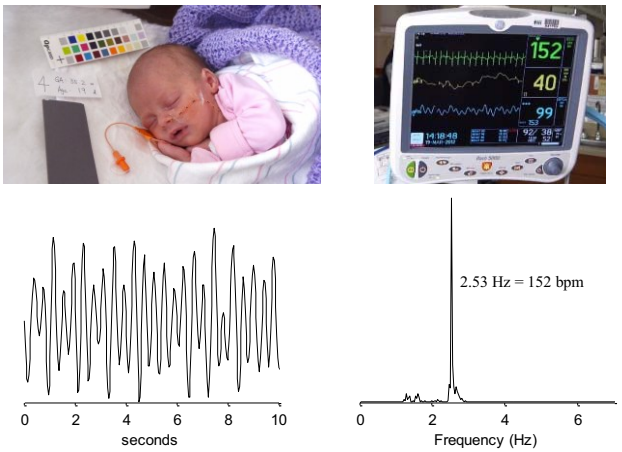

Figure 8: Results from video of a sleeping newborn. The actual heart rate is about 152 bpm (top right). Our method produces a clean signal (bottom left) and a frequency closely matching the ground truth.

to the ECG distributions, indicating that we do capture a real physiological variability. For 16 of the 18 subjects, we found that there was not a statistically significant difference between the ECG and the motion beat intervals. It is worth noting that this is a stronger test than is required in most clinical contexts. Typically heart rate variability (HRV) is used to dichotomize patients into high and low risk groups, so the precise shape of the distribution is not relevant. The relevant test would be whether the distribution of motion-generated intervals yields the same set of high risk individuals as ECG generated intervals. Since all of our subjects were healthy volunteers, we were not able to perform this test.

Several factors affected our results. First, our camera has a sampling rate of 30Hz. ECG used for HRV analysis normally has a sampling rate of at least 128 Hz. Cubic interpolation of our signal only partially addresses this discrepancy. Second, extra variability might be introduced during the pulse transit time from the abdominal aorta to the head. In particular, arterial compliance and head mechanics could affect our results. Third, the variable and suboptimal lighting conditions can affect our feature tracking. We believe this to be the case for several of our videos. Finally, our videos were only a maximum of 90 seconds long. Normally, HRV measures are computed over many hours to obtain reliable estimates.

An important future direction is to develop approaches for moving subjects. This is complicated because, as our results show, even other involuntary head movements are quite large in relation to pulse motion. Clearly with larger motions such as talking, more sophisticated filtering and decomposition methods will be needed to isolate pulse.

In this work we considered the frequency and variability of the pulse signal. However, head movement can offer other information about the cardiac cycle. If head displacement is proportional to the force of blood being pumped by the heart, it may serve as a useful metric to estimate blood stroke volume and cardiac output. Additionally, the direction of the movement could reveal asymmetries in blood flow into or out of the head. This might be useful for the diagnosis of a stenosis, or blockage, of the carotid arteries.

Another future direction is to better assess the strengths and weaknesses of the color and motion pulse estimation methods. Our results suggest that neither method is strictly more robust than the other in the presence of noise. However, further work needs to be done with varying lighting, skin tones, and distance from the camera to form a complete picture. In addition, we need to understand how sensitive the methods are to voluntary motions like talking or typing. For many applications, this is a critical factor. A motion-based approach is certainly better when the face is not visible. Based on these ideas, we believe that a combination of the color and motion methods will likely prove more useful and robust than using either one independently.

6. Summary

We described a novel approach that offers a non-invasive, non-contact means of cardiac monitoring. Our method takes video as input and uses feature tracking to extract heart rate and beat measurements from the subtle head motion caused by the Newtonian reaction to the pumping of blood at each heartbeat. A combination of frequency filtering and PCA allows us to identify the component of motion corresponding to the pulse and we then extract peaks of the trajectory to identify individual beats. When evaluated on 18 subjects, our method produced virtually identical heart rates to an ECG and captured some characteristics of inter-beat variability.

References

[1] L. Adde et al. Early prediction of cerebral palsy by computer-based video analysis of general movements: a feasibility study. *Developmental Medicine & Child Neurology*, 52:773–778, 2010.

[2] G. Bonmassar et al. Motion and ballistocardiogram artifact removal for interleaved recording of eeg and eps during mri. *NeuroImage*, 16:1127–1141, 2001.

[3] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[4] S. Chen et al. Quantification and recognition of parkinsonian gait from monocular video imaging using kernelbased principal component analysis. *BioMedical Engineering OnLine*, 10, 2011.

[5] M. Delano. A long term wearable electrocardiogram (ecg) measurement system. *Master's Thesis, Massachusetts Institute of Technology*, 2012.

[6] M. Garbey et al. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *IEEE Trnas Biomed Eng*, 54(8):1418–1426, 2007.

[7] D. He. A continuous, wearable, and wireless heart monitor using head ballistocardiogram (bcg) and head electrocardiogram (ecg). *Conf Proc IEEE Eng Med Biol Soc. 2011*, 2011, 2011.

[8] O. Inan et al. Robust ballistocardiogram acquisition for home monitoring. *Physiological Measurement*, 30:169–185, 2009.

[9] K. Kim et al. A new method for unconstrained pulse arrival time measurement on a chair. *J. Biomed. Eng. Res.*, 27:83–88, 2006.

[10] C. Liu et al. Motion magnification. *ACM Trans. Graph.*, 24:519.

[11] P. Lynch. Human head anatomy with external and internal carotid arteries. `http://www.flickr.com/photos/patrlynch/450142019`, 2007.

[12] M. Pediaditis et al. Vision-based human motion analysis in epilepsy - methods and challenges. *Proceedings of IEEE ITAB*, pages 1–5, 2010.

[13] Philips. Philips vitals signs camera. `http://www.vitalsignscamera.com`, 2011.

[14] M. Poh et al. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10):10762–10774, 2010.

[15] I. Starr et al. Studies on the estimation of cardiac output in man, and of abnormalities in cardiac function, from the hearts recoil and the bloods impacts; the ballistocardiogram. *The American Journal of Physiology*, 127:1–28, 1939.

[16] Z. Syed et al. Computationally generated cardiac biomarkers for risk stratification after acute coronary syndrome. *Sci Transl Med*, 3(102):102ra95, 2011.

[17] K. Tan et al. Real-time vision based respiration monitoring system. *Proceedings of CSNDSP*, pages 770–774, 2010.

[18] W. Verkruysse et al. Remote plethysmographic imaging using ambient light. *Optics Express*, 16(26):21434–21445, 2008.

[19] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *Proceedings of IEEE Conference on Computer Vision and Patern Recognition*, pages 511–518, 2001.

[20] C. Wang et al. Vision analysis in detecting abnormal breathing activity in application to diagnosis of obstructive sleep apnoea. *Proceedings of IEEE Eng Med Biol Soc*, pages 4469–4473, 2006.

[21] F. Wieringa et al. Contactless multiple wavelength photoplethysmographic imaging: a first step toward spo2 camera technology. *Ann. Biomed. Eng.*, 33(8):1034–1041, 2005.

[22] H. Wu et al. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph. (Proceedings SIGGRAPH 2012)*, 31(4), 2012.