The Impact of Video Compression on Remote Cardiac Pulse Measurement Using Imaging Photoplethysmography

Daniel J. McDuff[1], Ethan B. Blackford[2], and Justin R. Estepp[3]

[1] Microsoft Research, Redmond, WA, USA. (danmcduff@microsoft.com)

[2] Ball Aerospace, Fairborn, OH, USA. (ethan.blackford.ctr@us.af.mil)

[3] 711th Human Performance Wing, US Air Force Research Laboratory, Wright-Patterson AFB, OH, USA.

(justin.estepp@us.af.mil)

Abstract— **Remote physiological measurement has great potential in healthcare and affective computing applications. Imaging photoplethysmography (iPPG) leverages digital cameras to recover the blood volume pulse from the human body. While the impact of video parameters such as resolution and frame rate on iPPG accuracy have been studied, there has not been a systematic analysis of video compression algorithms. We compared a set of commonly used video compression algorithms (x264 and x265) and varied the Constant Rate Factor (CRF) to measure pulse rate recovery for a range of bit rates (file sizes) and video qualities. We found that compression, even at a low CRF, degrades the blood volume pulse (BVP) signal-to-noise ratio considerably. However, the bit rate of a video can be substantially decreased (by a factor of over 1000) without destroying the BVP signal entirely. We found an approximately linear relationship between bit rate and BVP signal-to-noise ratio up to a CRF of 36. A faster decrease in SNR was observed for videos of the task involving larger head motions and the x265 algorithm appeared to work more effectively in these cases.**

I. INTRODUCTION

Remote measurement of physiological signals has a number of advantages over traditional contact methods. It allows the measurement of vital signals unobtrusively and concomitantly. In recent years, a number of approaches for imaging-based measurement of physiology using digital cameras have been proposed. Imaging photoplethysmography (iPPG) captures variations in light reflected from the body due to blood volume changes in microvascular tissue [1]. Verkruysse *et al.* [2] demonstrated that sub-pixel variations in color channel measurements from a digital single lens reflex (DSLR) camera, when aggregated, could be used to recover the blood volume pulse. Subsequently, researchers have shown that iPPG methods can allow accurate measurement of heart rate [3], heart rate variability [4], breathing rate [4], blood oxygenation [5] and pulse transit time [6]. McDuff *et al.* [7] provide a comprehensive survey of approaches to iPPG.

A number of parameters influence the accuracy of iPPG measurements. These include the imager quality [8], and the frame rate and resolution of the images [9]. Sun *et al.* [8] compared remote physiological measurement using a low cost webcam and a high-speed color CMOS and showed similar signals were captured from both cameras, further supporting that iPPG is a practical method for scalable applications
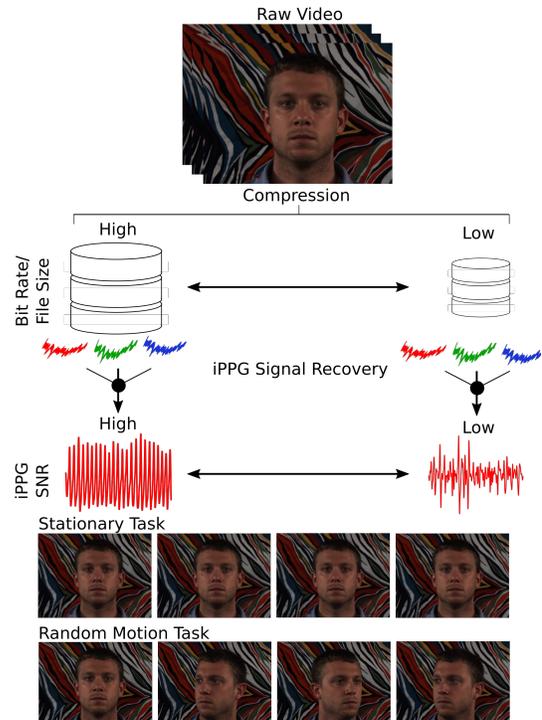
Fig. 1. We present a systematic analysis of the impact of video compression on physiological measurements via image photoplethysmography. We compare compression types and levels on over four hours of video from twenty-five participants performing stationary and random head motion tasks.

such as telemedicine. Blackford and Estepp [9] found that reducing frame rate from 120Hz to 30Hz and/or reducing image resolution from 658x492 pixels to 329x246 pixels had little impact on the accuracy of pulse rate measurements. Video compression is an important parameter that has not been systematically studied with regard to iPPG.

There are a number of methods for video compression that aim to reduce the bit rate whilst retaining the important information within a video. However, video compression algorithms are not designed with the intention of preserving photoplethysmographic data. On the contrary often compression algorithms make assumptions that small changes in pixel values between frames are not of high visual importance and discard them, influencing the underlying variations on which iPPG methods rely. We present a comparison of iPPG blood volume pulse (BVP) signal-to-noise ratios and pulse rate (PR) measurements from videos compressed using popular current

and next-generation codecs (x264 and x265).

In real-life applications motion tolerance of iPPG measurement is likely to be important. Previous work has proposed methods for reducing motion artifacts in iPPG measurements [10], [11], [12], [13], [14], [15], [16]. Due to the nature of inter-frame compression, compression is likely to have different impacts on physiological signal recovery depending on the level of head motion. Therefore, we evaluate results on videos for stationary and motion tasks.

Finding compression configurations that preserve valuable physiological data would allow new applications for iPPG measurement. For example, methods used for video recording/streaming through a web browser in video conferencing could be adapted to preserve iPPG data for analysis as part of a telehealth system. Additionally, alleviating the burden of storing raw video could enable sharing research datasets.

We analyzed a large dataset of uncompressed, raw videos with both stationary subjects and random head motions [13] in order to test the impact of video compression on the accuracy of remote physiological measurements. Participants (n=25) engaged in two, 5-minute tasks and were recorded using an array of cameras. Gold-standard electrocardiogram (ECG) measurements were captured alongside contact PPG measurements from the finger-tip. Figure 1 shows a summary of our study and examples of frames from the two tasks.

II. BACKGROUND

A. Video Compression

Raw video requires enormous amounts of storage; for example, each raw, uncompressed 5.5-minute video file collected in this study was 11.9 GiB in size. Collecting data from multiple cameras (9), numerous trials (12), and subjects (n=25) resulted in a total size of 31.50 TiB for the 247.5 hours of standard definition video. The inherently large size of raw video make its use infeasible outside of research and archival video storage. Herein we seek to better understand the trade offs related to iPPG derived measurements made from videos with varying levels of compression. The results of this evaluation will inform the final dataset which is planned to be made available to researchers working in this area.

Owing to the large file sizes imposed by raw video, compression is an essential element of almost every video system. Video encoding schemes typically employ similar methods in order to reduce the amount of data required to store or transmit video. Increasing the complexity of such methods improves coding efficiency and relies on advances in computational resources to allow the video to be decoded. These methods may be *lossless* and utilize principles of information theory to reduce data rates while allowing identical reproduction of the source video or *lossy* and discard less visually important information. Some of these methods include color space conversions, reducing inter-frame and intra-frame (motion) redundancy, and entropy coding of the data to produce an efficient binary representation.

Video is typically recorded from image sensors outfitted with a red, green, blue Bayer color filter array, where each pixel records light transmitted through a single filter. The resulting image may then be interpolated to derive a full color image with RGB values for each location, increasing the amount of data by a factor of three. Alternatively, images may be stored using alternate colorspaces, such as YUV. The YUV colorspace represents data as a single luma (Y), or achromatic brightness, component and two chrominance components, U/Cb and V/Cr representing blue-luma and red-luma differences, respectively. A popular use of the YUV colorspace, YUV420p, capitalizes on the visual system's greater perception of luminance differences over color or chrominance differences. Each image pixel location is represented by a luminance value while the chrominance values are subsampled every other row and column. As a result, a given block of 4 pixels requires 6 bytes of data (12bits/pixel) in the YUV420p colorspace rather than 12 bytes (24bits/pixel) in the RGB colorspace.

Intra-frame or image compression methods reduce spatial redundancy/correlation within the image. To do so, the image is subdivided into groups of pixels, sometimes referred to as macroblocks or coding units. Larger and more complicated block compositions can provide better visual appearance and greater coding efficiency at the expense of additional computational complexity. The blocks are then transformed from the spatial domain, often using the discrete cosine transform (DCT). The DCT is then divided by a quantization matrix and rounded. This process eliminates smaller coefficients and greatly reduces the number of values required to express the image. This process is also used in JPEG image coding.

Similarly, inter-frame compression reduces temporal redundancy/correlation between successive images in a group of pictures (GOP). A reference, I-frame (intra-) is encoded independently using intra-frame compression, as described above, and requires the most data to express. Between I-frames are P-frames (predicted), and B-frames (bi-directionally predicted). P-Frames require less data to express and consist of motion vectors describing changes from previous I- or P-frames. B-frames require the least amount of data and describe motion vectors using both past and future frames. The difference, or residual, between predicted frames and original frames are then transformed and quantized. This process reduces the coding for regions of little to no change between frames.

More advanced encoding schemes rely on more sophisticated techniques for determining I-, P-, and B-frames, their order and frequency for more efficient encoding via greater complexity. Similarly, adaptive quantizers, such as a constant rate factor (CRF) may be used to maintain video quality across frames with different amounts of motion content and image complexity. Finally, inter- and intra- frame compression are utilized in tandem for even greater efficiency.

B. Consequences of Compression for iPPG Signal Recovery

Various stages of video compression are likely to have detrimental effects on iPPG measurements. For example, the chroma subsampling to the YUV colorspace reduces the number of samples used to represent chromaticity, where the BVP predominately resides [17]. Furthermore, the BVP is an imperceptible color change that often occurs at sub-

noise levels for individual pixels and may be only measured after averaging over a region. An otherwise imperceptible change in the video, such as this, could easily be deteriorated or eliminated by spatial or motion compression. That is, relatively spatially-homogeneous regions may be quantized jointly. This will limit the subsequent benefits derived from averaging across regions, as frequently performed in iPPG signal processing. Additionally, the time-evolution of the BVP may be quantized and discarded as it is an otherwise imperceptible change between video frames. Given these potential effects, it is important to understand the potential consequence of video compression on iPPG measurements.

C. Video Compression Standards

Numerous and varied codecs exist implementing variations of the methods described above. Some of the most popular have been developed as standards of the Moving Picture Experts Group (MPEG) and International Telecommunication Union Telecommunication Standardization Sector (ITU-T). These include MPEG-2 Part 2/H.262, utilized for DVD video and over-the-air, digital-TV broadcasts [18]. MPEG-4 Part 10, AVC/H.264 a current-generation standard, utilized by Blu-ray video and various high definition (HD) video providers, which achieved an approximate doubling in encoding efficiency over H.262 and enabled network video streaming [19], [20]. Finally, next-generation standards such as MPEG-H Part 2, HEVC/H.265 utilize increasingly complex encoding strategies for an approximate doubling in encoding efficiency over H.264 [21], [22]. H.265 is designed to handle ultra-high definition video and provide improved mobile and Internet video streaming. Other next-generation standards, have similar efficiency goals and employ analogous techniques to achieve them. Other standards vying for the role of leading next generation codec include VP9 and its successor AV1, both open source and royalty free.

Modern video coding standards typically represent a compressed video format and the necessary operations to decode a compliant video. This ensures a uniform output from a given video while providing flexibility in the operations and compression methods used. Various frameworks may then be used to generate compressed videos.

D. Codec Evaluations

Various groups have published systematic evaluations of video codecs within the three-dimensional trade-off space of encoding efficiency, video quality, and computational complexity. Encoding efficiency is typically assessed using average bit rate. Computational complexity is typically assessed using a benchmark processing time. The gold standard for video quality is human subjective perception assessed by a Mean Opinion Score (MOS). Alternatively, video quality may be assessed using one of numerous automated metrics including Peak Signal-to-Noise Ratio (PSNR), Multi-Scale Structural Similarity (MS-SIM), and PSNR with contrast Masking of DCT basis functions based on the Human Visual System (PSNR-HVS-M). PSNR-HVS-M adapts PSNR measurements with a contrast sensitivity function which takes
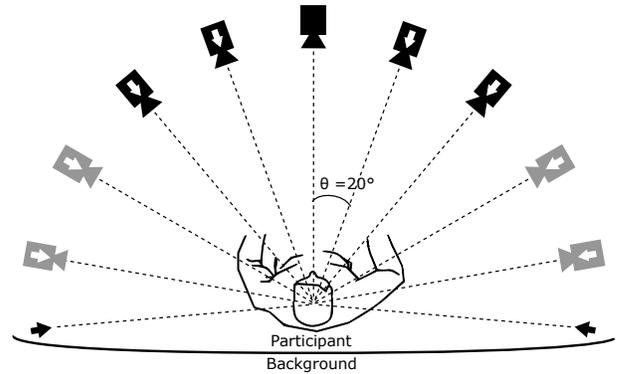

Fig. 2. Imager array and lighting setup used for data collection. Arrows indicate light sources. Cameras were positioned at the subject's eye level. Recordings were collected with nine cameras; however, previous work showed BVP recovery with the central five cameras was comparable, or better, to that with all nine [13]. To reduce excess computation only images from the central five cameras (shown in black) were used in this analysis.

into account the maximal non-perceptible distortion between DCT coefficients for the macroblocks of an image. This adaptation allows PSNR-HVS-M to correlate well with MOS values [23]. PSNR-HVS-M was evaluated in this work using the EPFL Video Quality Metric Tool (VQMT) [24].

Evaluations conducted by Moscow State University [25], [26] and Netflix [27] assessing these tradeoffs have shown the strengths of x264 as a current-generation codec and x265 as a next-generation codec. These evaluations helped guide our methods while particularly focusing on effects of compression on physiologically relevant signals derived from iPPG.

In this work, open-source tools including FFmpeg and codecs compliant with H.264 (x264) [28] and H.265 (x265) [29] were used. Particular attention was paid to maintaining video compatibility with MATLAB (R2016b on Windows 10), due to the popularity of the MATLAB environment for data and video processing in this area of research. The resulting encoded videos were assessed on basis of average bit rate of the video, video quality assessed by PSNR-HVS-M, pulse rate measurement accuracy, and BVP signal-to-noise ratio.

III. EXPERIMENT

A. Apparatus

Imager Array: A multi-imager array was used to collect synchronized videos of the participant. Nine cameras spaced every $20°$ across a semicircular arc (total of $160°$) were used to collect videos. The video from the center most five cameras was analyzed based on the results of [13]. The cameras were positioned at the participant's eye level. Figure 2 shows the arrangement of the cameras. Synchronization of the image capture from the array was achieved using a PCIe-6323 data acquisition card (National Instruments, Austin, Texas, USA) to generate a hardware-timed (120 Hz, 90% duty cycle) trigger/exposure control signal, resulting in a capture rate of 120 (fps) and 7.5 ms exposures.

Imagers: Scout scA640-120gc (Basler, Ahrensburg, Germany) GigE-standard, color, progressive-scan, CCD cameras with external triggering and exposure control, capturing 8-bit,

658x492 pixel, raw BG Bayer-format images were used for video capture. The imagers were equipped with 16 mm fixed focal length lenses (HF16HA-1B, Fujinion, Tokyo, Japan) to maximize the area of the participant's face in each frame.

Lighting: Experimental lighting was provided by 10 SoLux, full color spectrum (5000K) bulbs (Solux MR-16, Tailored Lighting, Inc., Rochester, New York, USA) equipped with a frosted diffuser. The bulbs were mounted 0.5 m above the imagers and positioned as shown in Figure 2.

B. Contact Physiological Measurements

Gold-standard physiological measurements were collected and natively synchronized with the camera exposure/acquisition trigger signal. PPG and ECG signals were measured using a research-grade, biopotential acquisition unit with peripheral-physiological and trigger signal inputs (ActiveTwo, BioSemi B.V., Amsterdam, The Netherlands). ECG was measured from the chest using three leads. BVP was calculated via the PPG signal from the index fingertip on the left hand (MLT1020FC IR Plethysmograph Finger Clip, ADInstruments, Inc., Colorado Springs, Colorado, USA).

C. Participants

Twenty-five participants (17 male, 18 to 28 years, mean age 23.7 years) were recruited to take part in our experiment. The experimental protocol was reviewed and approved by the Air Force Research Laboratory Institutional Review Board and performed in accordance with all relevant institutional and national guidelines and regulations. All prospective participants received a study briefing and completed comprehensive written informed consent prior to their voluntary participation. Participants were compensated for their time. Nine individuals were wearing glasses and eight had facial hair.

D. Tasks

Participants completed six, 5-minute tasks in front of two different background screens resulting in one hour of video per participant (6 x 5 x 2 = 60 minutes). Two of these tasks are analyzed in this evaluation. We did not analyze all six tasks due to the extremely time consuming process of producing compressed versions of all the videos with many CRF levels. The two tasks are representative of the data overall.

Stationary Task: Participants were asked to look forward and remain still throughout the task.

Random Motion Task: Participants were asked to reorient their head position once per second to a randomly chosen imager in the array. Thus simulating random head motion and imposing additional noise near typical, resting pulse-rate frequencies (~60 bpm). The random sequence was provided to the participant via a pre-generated audio recording.

IV. VIDEO COMPRESSION ANALYSIS

The face videos were originally recorded in lossless, raw image format. We tested a set of commonly used lossy compression codecs on these videos in order to quantify the impact of compression on physiological parameter estimates. Compression was performed using the latest FFmpeg
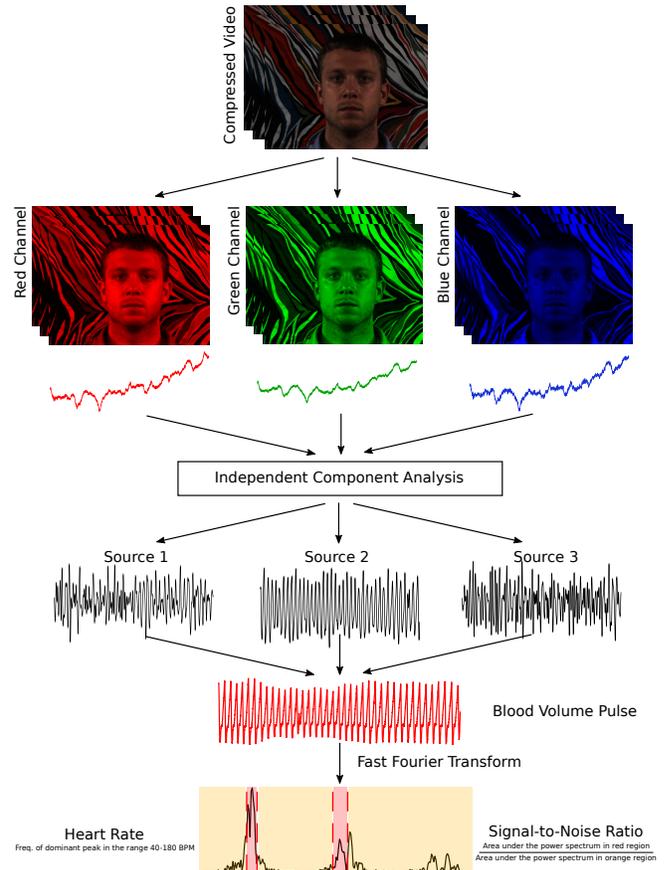

Fig. 3. Method used for recovering the blood volume pulse from the video sequences. Spatial averages of the red, green, and blue channel frames are calculated to form three observation signals. Independent component analysis is used to recover three underlying source signals. The source signal with the greatest peak energy in the frequency domain is selected as the blood volume pulse estimate. The Fast Fourier Transform of the BVP signal is used to calculate the pulse rate and signal-to-noise ratio.

Windows 64-bit binary release (at time of testing: N-82324-g872b358) [30]. FFmpeg provides an extensive command line interface for video transcoding, filtering, and streaming of video, images, and audio. FFmpeg is a versatile tool supporting a great variety of formats and containers. An example of the command structure used is shown below.

```
ffmpeg -i [input] -c:v [codec] -crf [x] ...
    -pix_fmt yuv420p [output.mp4]
```

A. Codecs

x264: An open-source, current-generation codec producing H.264 compliant videos [28].

x265: An open-source, next-generation codec built on the x264 code base producing H.265 compliant videos [29].

We use the x264 and x265 codecs released under the terms of the GNU General Public License and contained in the FFmpeg release above [30].

B. Parameters

We chose to vary the Constant Rate Factor (CRF) for both compression types in order to understand the trade-offs between bit rate, video quality, and accuracy of pulse recovery. CRF values control the adaptive quantization parameter to

| | Stationary Task | | | | | | Random Motion Task Task | | | | | |
| | x264 | | | x265 | | | x264 | | | x265 | | |
CRF	SNR (dB)	PR Err. (BPM)	Bit Rate (kb/s)	SNR (dB)	PR Err. (BPM)	Bit Rate (kb/s)	SNR (dB)	PR Err. (BPM)	Bit Rate (kb/s)	SNR (dB)	PR Err. (BPM)	Bit Rate (kb/s)
Cont.	-4.31	0.57	N/A	-4.31	0.57	N/A	-2.99	0.40	N/A	-2.99	0.40	N/A
Raw	-7.80	1.78	$3.1*10^5$	-7.80	1.78	$3.1*10^5$	-8.70	1.18	$3.1*10^5$	-8.70	1.18	$3.1*10^5$
1	-8.50	1.71	$2.0*10^4$	-9.66	2.65	$3.8*10^4$	-11.3	4.02	$2.4*10^4$	-9.66	4.26	$4.3*10^4$
3	-9.03	2.11	$1.2*10^4$	-9.88	3.89	$2.4*10^4$	-11.8	5.88	$1.5*10^4$	-11.2	4.27	$2.8*10^4$
6	-9.68	2.17	$5.5*10^3$	-10.1	3.32	$1.1*10^4$	-12.6	7.57	$7.0*10^3$	-11.2	4.62	$1.3*10^4$
9	-10.4	2.68	$2.3*10^3$	-11.8	5.94	$4.2*10^3$	-12.6	6.51	$3.2*10^3$	-12.0	6.34	$5.5*10^3$
12	-10.9	2.31	$8.9*10^2$	-12.6	7.25	$1.6*10^3$	-12.6	6.85	$1.5*10^3$	-12.6	6.85	$2.5*10^3$
15	-10.8	3.24	$3.6*10^2$	-13.1	7.84	$7.5*10^2$	-13.0	8.00	$7.8*10^2$	-12.7	6.85	$1.3*10^3$
18	-11.3	3.35	$2.0*10^2$	-13.9	7.49	$4.1*10^2$	-13.0	8.66	$4.8*10^2$	-13.2	8.26	$7.8*10^2$
24	-12.3	4.38	$1.1*10^2$	-13.9	9.51	$1.7*10^2$	-13.5	8.31	$2.3*10^2$	-13.9	9.90	$3.6*10^2$
30	-13.1	7.01	$7.8*10^1$	-13.9	10.95	$8.9*10^1$	-14.0	10.6	$1.3*10^2$	-14.1	9.48	$1.9*10^2$
36	-13.3	9.93	$6.7*10^1$	-13.7	8.87	$5.8*10^1$	-14.1	10.5	$8.8*10^1$	-14.4	9.29	$1.0*10^2$
42	-13.6	10.7	$5.0*10^1$	-13.8	9.29	$4.1*10^1$	-14.0	11.2	$5.7*10^1$	-14.2	11.5	$6.3*10^1$
48	-13.4	10.2	$4.8*10^1$	-13.8	11.8	$3.3*10^1$	-14.1	9.97	$5.2*10^1$	-13.7	10.5	$4.5*10^1$
51	-13.5	9.31	$8.9*10^1$	-13.8	9.85	$3.4*10^1$	-14.0	10.3	$9.4*10^1$	-14.1	12.1	$4.6*10^1$
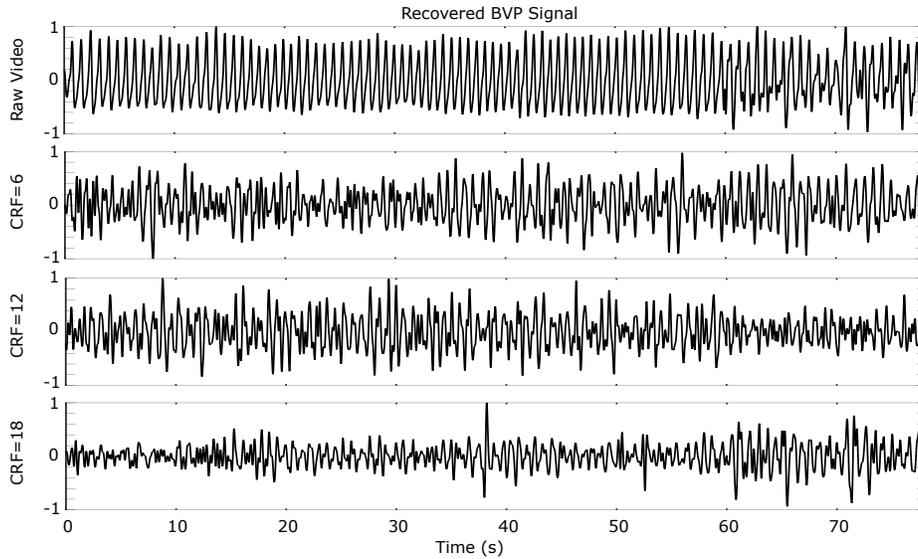
Fig. 4. Examples of the recovered blood volume pulse waves from the raw and compressed videos using the x264 method and CRF = 6, 12, and 18.

provide constant video quality across frames of varying complexity and motion. CRF values range from 0 to 51, with 0 being lossless quality but requiring the greatest average bit rate and 51 being the most lossy but producing the least average bit rate. Decreasing the CRF value by six results in an approximate doubling in average bit rate. For most videos produced for strictly visual purposes, typically CRF values are set in a range between 18 and 28. The implementations of CRF in x264 and x265 are similar but not identical. For completeness and anticipating effects on iPPG related signals prior to degredations in visual quality, we varied CRF in the range: $\{1, 3, 6, 9, 12, 15, 18, 24, 30, 36, 42, 48, 51\}$ for both codecs and compare with uncompressed (raw) videos.

V. RECOVERY OF THE BLOOD VOLUME PULSE

The BVP signal was recovered from the video sequences using a blind source separation approach. A whole-frame, spatial average of the color channel pixel values in each frame was calculated to form time-varying observation signals. The resulting signals, $r_1(t)$, $g_1(t)$, $b_1(t)..$, $r_5(t)$, $g_5(t)$, $b_5(t)$, represent the amplitudes of the recorded RGB signals from the five cameras at time point, t.

For the analysis, the 5-minute videos were broken into five, one-minute sections. The observations were detrended using a smoothness priors approach (λ=1000) [31]. Independent component analysis (ICA), a blind source separation technique, was used to recover a set of three source signals from the observations by maximizing the non-gaussianity of the signals. The JADE ICA implementation was used [32].

The resulting source signals were filtered using a bandpass filter with low and high frequency cut-offs at 0.75 Hz (45 BPM) and 3 Hz (180 BPM) respectively. The dominant iPPG signal was selected from the set of source signals based on the concentration of power in the frequency domain using the method proposed in [33]. The resulting signal was used as the BVP estimate. Figure 3 shows a summary of our approach. Figure 4 shows some qualitative examples of the recovered BVP waves using different CRFs. It is clear how the BVP wave becomes considerably degraded with greater compression. Pulse rate variability estimates on the compressed videos would be much less accurate due to the absence of clear peaks.
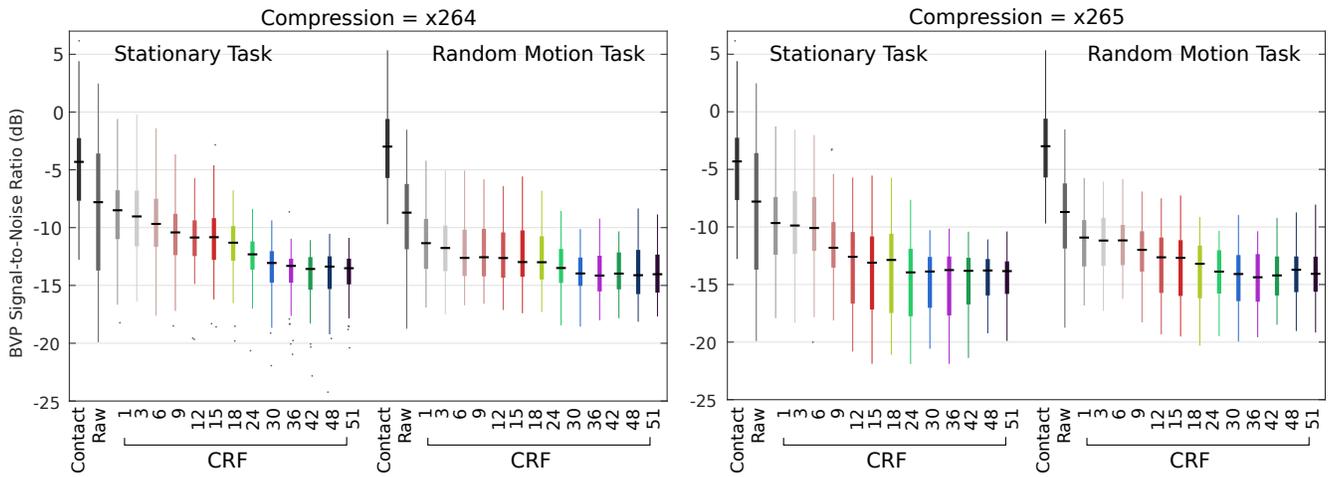
VI. RESULTS AND DISCUSSION

Fig. 5. Box plots of the blood volume pulse (BVP) signal-to-noise ratio for the x264 (left) and x265 (right) compression methods. The results for each task and constant rate factor (CRF) are shown in comparison to the results from a contact BVP sensor and the raw video. The gold-standard pulse rate frequency, used to calculate the SNR, was measured from the ECG signal.
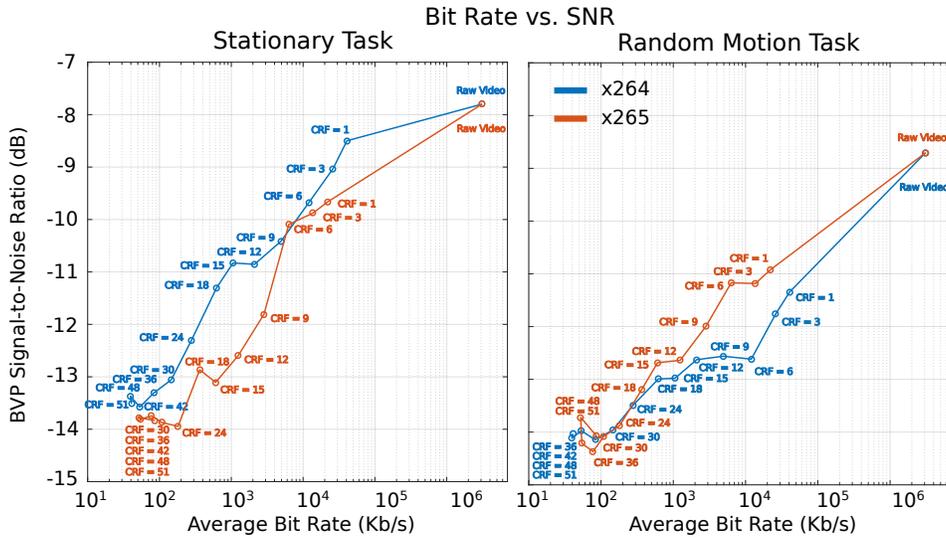

Fig. 6. Plots of the average bit rate (log-scale axis) versus the BVP SNR (in dB) for different compression CRF values. The results for the x264 (blue) and x265 (red) codecs and the two tasks, stationary (left) and random motion (right), are shown. As both scales are logarithmic, there is a linear relationship between bit rate and SNR. For the stationary task, the x264 codec performed better overall. For the motion task, the x265 codec performed better.
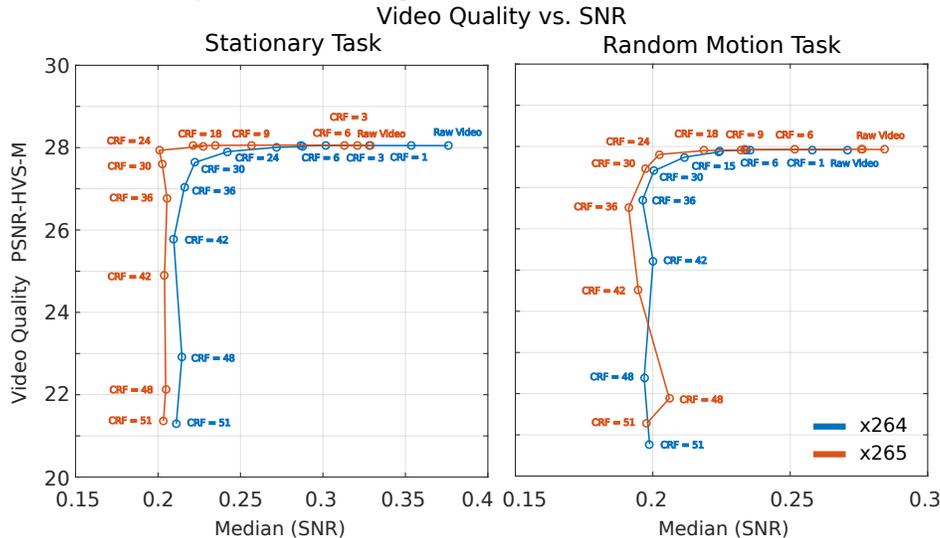

Fig. 7. Plots of the average video quality PSNR-HVS-M (in dB) versus the BVP SNR (in dB) for different compression CRF values. The results for the x264 (blue) and x265 (red) codecs and the two tasks, stationary (left) and random motion (right), are shown.

To assess the impact of video compression on the performance of the iPPG measurements, we calculate the absolute error in pulse rate compared with measurement from the ECG signal and the signal-to-noise ratio (SNR) in the recovered blood volume pulse signal. The pulse rate error and SNR were calculated for each one-minute non-overlapping window from each session (excluding the first and last 15 seconds of the session) for total of 125 minutes per task.

The pulse rate was calculated as $60/\overline{IBI}$, where $\overline{IBI}$ is the mean of the inter-beat intervals. The SNR is calculated using the method described by de Haan and Jeanne [17], where the ratio of power within a template that includes the contact sensor heart rate frequency and the first harmonic frequency is divided by the power outside of the template region (from 0 to 240 BPM). This is represented in Figure 3.

Table I shows the numerical results for median SNR and pulse rate error alongside the median bit rate for each compression case. Figure 5 shows box plots of the BVP SNR for the x264 (left) and x265 (right) compression algorithms. In each plot we show the results for the stationary and random motion tasks. For the uniform motion tasks, the SNR for the raw video measurements was lower than for the contact BVP measurements but the recovered waveform was still very good. The median pulse rate error was 0.57 (contact sensor) and 1.78 (raw images). As expected, increasing the CRF steadily reduces the SNR of the resulting PPG measurement. Above a CRF of 30, the SNR does not continue to decrease, suggesting that the signal has been degraded sufficiently that there is little physiological information remaining. This result is also reflected in the pulse rate error that is consistently around 10 BPM for CRF values below 30.

Figure 6 shows the trade-off between video bit rate and the PPG SNR. Again, we show the results for the x264 (blue line) and x265 (red line) compression algorithms and the stationary (left) and random motion (right) tasks. Compression allows the video bit rate to be reduced significantly. Videos compressed at a CRF of 6 have a bit rate 100 times smaller than the raw video. However, results clearly illustrate that at CRF values of 30 or above, the signal is severely reduced. This is worsened in the presence of additional noise sources such as large head movements. Video encoding is computationally intensive, but may be performed more quickly and efficiently directly on hardware. In many cases, constraints on bandwidth and/or storage space will dictate whether compression is required, in which case the reduction in signal will be accepted.

We also plot the relationship between the video quality (as measured by the PSNR-HVS-M metric) and the BVP signal-to-noise ratio. Figure 7 displays these results. The two values are not highly correlated and the results provide evidence that the degradation of the BVP signal begins before there are deteriorations in image quality. Once image quality begins to decrease rapidly (CRF > 36) the BVP signal has moreorless been lost. In this case, BVP SNR is a preferred metric to those assessing video quality.

It may also be worth considering that, when compressing videos, it is often preferred to reduce the bit rate by down-sampling the video resolution prior to encoding rather than relying solely on significant amounts of lossy compression. Visual artifacts related to upscaling the video after decoding are perceived less harshly than strong compression artifacts. Regarding videos collected for iPPG, a similar strategy may be beneficial. Previously it was shown that large reductions of either video resolution or frame rate did not strongly impact iPPG pulse rate measurements [9]. A balanced approach may yield better results than relying on one method alone to reduce the video bit rate to acceptable levels.

VII. CONCLUSIONS

Remote physiological measurement, in particular imaging photoplethysmography (iPPG), has received a lot of attention in recent years due to the great potential of low-cost measurement of vital signs. Compression methods enable reduced video bit-rates whilst preserving visual content. This enables web streaming and other bandwidth limited applications; however, compression algorithms were not designed with iPPG in mind. We performed a systematic analysis of compression methods and parameters to evaluate the impact on remote physiological measurement.

We tested popular compression methods (x264 and x265) and varied the constant rate factor to obtain a range of average bit rates and video qualities. Our results suggest a considerable drop in SNR between raw and compressed videos (even before visual quality becomes noticeably degraded). As the compression constant rate factor is increased, the bit rate and SNR drop linearly. Videos with a bit rate of 10Mb/s still retained a BVP with reasonable SNR and the pulse rate estimation error was 2.17 BPM. The results suggest the x265 compression method may be more effective than x264 on videos featuring greater motion.

REFERENCES

[1] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, p. R1, 2007.

[2] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Optics express*, vol. 16, no. 26, pp. 21 434–21 445, 2008.

[3] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Optics Express*, vol. 18, no. 10, pp. 10 762–10 774, 2010.

[4] ——, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 7–11, 2011.

[5] A. R. Guazzi, M. Villarroel, J. Jorge, J. Daly, M. C. Frise, P. A. Robbins, and L. Tarassenko, "Non-contact measurement of oxygen saturation with an rgb camera," *Biomedical optics express*, vol. 6, no. 9, pp. 3320–3338, 2015.

[6] D. Shao, Y. Yang, C. Liu, F. Tsow, H. Yu, and N. Tao, "Noncontact monitoring breathing pattern, exhalation flow rate and pulse transit time." *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 11, p. 2760, 2014.

[7] D. J. McDuff, J. R. Estepp, A. M. Piasecki, and E. B. Blackford, "A survey of remote optical photoplethysmographic imaging methods," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 6398–6404.

[8] Y. Sun, C. Papin, V. Azorin-Peris, R. Kalawsky, S. Greenwald, and S. Hu, "Use of ambient light in remote photoplethysmographic systems: comparison between a high-performance camera and a low-cost webcam," *Journal of biomedical optics*, vol. 17, no. 3, pp. 0 370 051–03 700 510, 2012.

[9] E. Blackford and J. Estepp, "Effects of frame rate and image resolution on pulse rate measured using multiple camera imaging photoplethysmography," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2015, pp. 94 172D–94 172D.

[10] G. Cennini, J. Arguel, K. Akşit, and A. van Leest, "Heart rate monitoring via remote photoplethysmography with motion artifacts reduction," *Optics express*, vol. 18, no. 5, pp. 4867–4875, 2010.

[11] A. Chung, X. Y. Wang, R. Amelard, C. Scharfenberger, J. Leong, J. Kulinski, A. Wong, and D. A. Clausi, "High-resolution motion-compensated imaging photoplethysmography for remote heart rate monitoring," in *SPIE BiOS*. International Society for Optics and Photonics, 2015, pp. 93 160A–93 160A.

[12] G. de Haan and A. van Leest, "Improved motion robustness of remote-ppg by using the blood volume pulse signature," *Physiological measurement*, vol. 35, no. 9, p. 1913, 2014.

[13] J. R. Estepp, E. B. Blackford, and C. M. Meier, "Recovering pulse rate during motion artifact with a multi-imager array for non-contact imaging photoplethysmography," in *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1462–1469.

[14] Y. Sun, S. Hu, V. Azorin-Peris, S. Greenwald, J. Chambers, and Y. Zhu, "Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise," *Journal of biomedical optics*, vol. 16, no. 7, pp. 077 010–077 010, 2011.

[15] M. van Gastel, S. Stuijk, and G. de Haan, "Motion robust remote-ppg in infrared," 2015.

[16] W. Wang, S. Stuijk, and G. de Haan, "Exploiting spatial redundancy of image sensor for motion robust rppg." *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 415–425, 2015.

[17] G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rppg," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886, 2013.

[18] ITU-T, "Rec. H.262 - Information technology - Generic coding of moving pictures and associated audio information: Video," International Telecommunication Union Telecommunication Standardization Sector (ITU-T), Tech. Rep., 1995.

[19] ——, "Rec. H.264 - Advanced video coding for generic audiovisual services," International Telecommunication Union Telecommunication Standardization Sector (ITU-T), Tech. Rep., 2003.

[20] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H. 264/AVC video coding standard," *IEEE Transactions on circuits and systems for video technology*, vol. 13, no. 7, pp. 560–576, 2003.

[21] ITU-T, "Rec. H.265 - High efficiency video coding," International Telecommunication Union Telecommunication Standardization Sector (ITU-T), Tech. Rep., 2013.

[22] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[23] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, "On between-coefficient contrast masking of DCT basis functions," in *Proceedings of the third international workshop on video processing and quality metrics*, vol. 4, 2007.

[24] P. Hanhart. Video Quality Measurement Tool (VQMT) 1.1. Ecole Polytechnique Federale de Lausanne (EPFL), Multimedia Signal Processing Group (MMSPG). [Online]. Available: http://mmspg.epfl.ch/vqmt/

[25] D. Vatolin, D. Kulikov, and M. Arsaev. Moscow State University 2012 MPEG-4 AVC/H.264 video codecs comparison. [Online]. Available: http://www.compression.ru/video/codec_comparison/h264_2012/

[26] D. Vatolin, D. Kulikov, M. Erofeev, S. Dolganov, and S. Zvezdakov. Moscow State University 2016 HEVC video codecs comparison. [Online]. Available: http://compression.ru/video/codec_comparison/hevc_2016/

[27] J. De Cock, A. Mavlankar, A. Moorthy, and A. Aaron, "A large-scale video codec comparison of x264, x265 and libvpx for practical vod applications," in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2016, pp. 997 116–997 116.

[28] VideoLAN Organization. x264 home page. [Online]. Available: http://www.videolan.org/developers/x264.html

[29] MulticoreWare Inc. x265 home page. [Online]. Available: http://x265.org/

[30] F. Bellard, M. Niedermayer *et al.* Ffmpeg. [Online]. Available: http://ffmpeg.org

[31] M. P. Tarvainen, P. O. Ranta-Aho, P. A. Karjalainen *et al.*, "An advanced detrending method with application to hrv analysis," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 2, pp. 172–175, 2002.

[32] J.-F. Cardoso, "Multidimensional independent component analysis," in *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, vol. 4. IEEE, 1998, pp. 1941–1944.

[33] D. McDuff, S. Gontarek, and R. Picard, "Improvements in remote cardio-pulmonary measurement using a five band digital camera." *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 10, pp. 2593–2601, 2014.