Robust Contact-Free Heart Rate Measurement From Human Face Videos Using a Webcam

Muhammad Waqar, Reyer Zwiggelaar, and Bernard Tiddeman

***Abstract*—Heart rate measurement plays an important role in human health assessment and there have been a number of methods suggested in order to monitor it remotely with more ease and comfort. Contact-free heart rate measurement is one way to make it user-friendly and can also be used for covert surveillance. The performance of previously developed methods is found to be efficient and accurate, in controlled conditions. In realistic and more challenging scenarios, their performance is degraded, as each method has its limitations. Typically, the performance is dependent on controlled lighting and limited subject movement. More realistic situations require more robust contact-free ways to measure the heart rate. In this paper, we have proposed a method that can overcome many of the problems of light reflection and subject's movements while measuring the heart rate from human faces remotely using an ordinary webcam. We find the robust mean of the skin pixel's color values and calculate the least-squares error optimal filter using our training data set to estimate the heart rate and obtain a cleaner pulse signal. This technique has potential for advancing human biometric systems employing the pulse signal obtained. Results obtained show that our method has significantly improved an existing methods.**

***Index Terms*—Fast Fourier transforms, Independent component analysis, photoplenthysmography (PPG).**

I. Introduction

HEART rate refers to the number of times a human heart beats in one minute. The normal resting heart rate of a healthy person ranges between 60 and 80 beats per minute (BPM) [1]. The continuous monitoring of the heart rate is very important to assess human health. Several different methods are currently in practice to measure the heart rate such as electrocardiogram (ECG), pulse oximetry and sphygmology. ECG (sensors attached to subject's body detects electrical activity of the heart every time it beats) is consider to be the gold standard for accurately measuring the heart rate, but the need to wear a chest band and associated cables makes it quite uncomfortable and restricts movement. Pulse oxymetry (heart rate measurement using pulse oximeter) and sphymology (pulse counting by placing finger over subject's wrist) are relatively comfortable but they are not very accurate in some situations such as cold hands and illness causes error. Contact-free heart measurement is therefore very demanding and hence it has been an interesting area for researchers to work on over the past years [6][7][8].

M. Waqar is with the Department of Computer Science, Aberystwyth University, Aberystwyth, Wales, UK (e-mail: muw1@aber.ac.uk).

R. Zwiggelaar is with the Department of Computer Science, Aberystwyth University, Aberystwyth, Wales, UK (e-mail: rrz@aber.ac.uk)

B. Tiddeman is with the Department of Computer Science, Aberystwyth University, Aberystwyth, Wales, UK (e-mail: bpt@aber.ac.uk).

Manuscript received October 20, 2016

Non-contact heart rate measurement was first introduced by Pavlidis *et al.*, in 2007 [2]. They employed a thermal camera as a sensing element to capture facial images and based on the thermal signal information emitted from facial vessels, measured the heart rate [2][3]. Gatto also used a thermal camera to measure the heart rate by monitoring the flow of blood on the forehead area and analyzed the fluctuations in the thermal energy released with respect to every heartbeat [4]. Takano and Ohta measured the pulse rate using time-lapse images captured by means of a CCD camera. They analyzed the variations in the average brightness of a participant's facial skin images [5].

Since CCD and thermal cameras can be very expensive, in the past few years researchers have reported on the measurement of the heart rate using inexpensive computer webcams. Poh *et al.*, in 2010 introduced a non-contact, automated heart rate measurement method using an ordinary webcam. They tracked human faces from color videos and extracted the average color of a region of interest (ROI) for each frame. They then applied blind source separation (BSS) on the resulting signals to decompose them into independent components and finally applied the Fast Fourier Transform (FFT) to the pulse component to estimate the heart rate [6]. Wu *et al.*, presented a method called Eulerian Video Magnification to measure the heart rate remotely. This method treated a video sequence as an input, spatially decomposed it and applied temporal filtering on the images. The expected signal was then amplified to visualize the flow of blood in the face and to reveal subtle motions [7]. De haan and Jeanne highlighted the problem of motion noise while measuring the heart rate and presented a chrominance-based remote pulse rate method to overcome such problems [8].

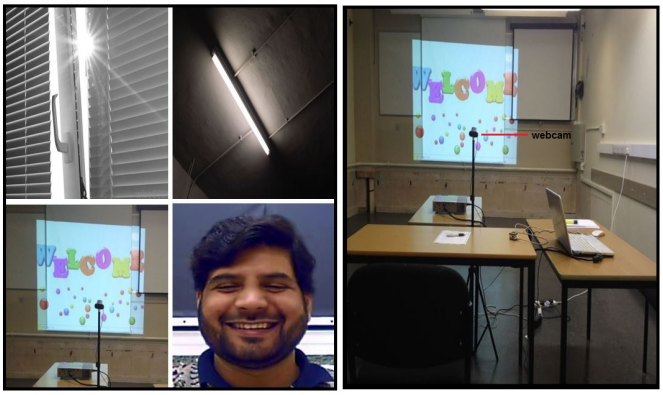
The main idea of the previous work related to cardiac pulse measurement from human faces is based on the phenomenon of Photoplenthysmography (PPG), which states that, when a heart beats, it produces a pulse wave that travels through the whole body's vascular system. When this wave reaches the face, it leaves a small change in the volume of blood, which causes the skin color to change [9]. The naked eye cannot view such a subtle changes but we can measure the intensity of the absorbed light by means of images taken from an ordinary webcam. The previous methods performed well, but only with controlled lighting and little motion, and their efficiency degraded as they were exposed to more realistic and challenging scenarios. The two main problems that affect the accuracy of existing methods includes illumination reflections and movements of the participant, which we have addressed in our work and have proposed a solution to overcome them.

We present a robust mean algorithm to employ on the skin color pixels of the ROI (cheeks) and calculate the least-squares error optimal filter using our dataset to measure a more accurate heart rate and obtain a clean pulse signal. We have trained and tested our system on a relatively large data set of 40 participants in more challenging environments and compared our results with the ground truth heart rate obtained by means of a pulse oximeter. We have obtained promising results while drawing a comparison of our method with other state of the art methods.

II. PROPOSED METHOD

A. Experimental Setup

Our dataset consists of 40 participants (age 18-40) of both genders from different origins (Asians, Africans and Caucasians) and skin colors. Informed consent was obtained from all the participants before the experiment. All experiments were conducted indoors with office florescent lights (one light was flickering throughout the experiments) and indirect sunlight as the main sources of illumination and a projector which reflected light making the lighting conditions more challenging. An ordinary Logitech webcam (C310) was used in our experiments which has the following specifications: 1.3 megapixels (mp) video capture capability, 24-bit RGB with 3 channels (8 bits per channel), 15 frames per second with a pixel resolution of 640 x 480. We asked the participant to sit and relax on a chair in front of a webcam at a distance of approximately 0.5 meter. We have then played a video clip on the projector screen that contains different emotional cues (happy, sad and scared videos) and asked the participant to view it while we have recorded a color video of the participant's face. The purpose of showing them the video clip is to add projector light to our lighting conditions and to encourage the participants to move their facial features (eyes, nose, mouth, head etc.) to a certain degree. The experimental scenario was hence made more realistic and challenging as we included extra lighting and movement artifacts. Fig. 1(a) presents the capturing conditions. We have recorded a color video of each participant for the duration of one minute synchronized with the projector display. All recorded videos were saved in .AVI format on a laptop under unique identifiers. We used a pulse oximeter (Contec CMS50E) in order to record the ground truth heart rate simultaneously for the training and testing purposes. Fig. 1(b) shows the experimental setup.



(a) (b)

Fig. 1. (a) Challenging capture parameters includes indirect sunlight, light flickering, projector light reflection, and participant's facial features movement.(b) Experimental setup.

B. Face Tracking and ROI Extraction

The color of facial skin changes with respect to every heartbeat and the RGB color sensor of a webcam can detect subtle changes of blood volume in the facial area. We refer to this small reflected light color change as a plethysmographic (PPG) or pulse signal that can be used to estimate the heart rate.

Face tracking is an important component in order to get the heart rate measured accurately. In most of the previous work, the OpenCV face detection method, which is based on the work of Viola and Jones [10] was employed to track the faces from the recorded videos. The problems of false positives (yielding multiple faces), false negatives (no face detected), and wide variation in how tight fitting the detected face box is, limits the performance of this method and therefore result in loss of data frames containing the face. We have used Saragih's Face Tracker [11] to track the faces from the recorded videos because it ensures no loss of frames for the detected face. It detects and tracks 68 points on different features of the face (eyes, nose, cheeks and lips) [14]. Saragih et al. performed a qualitative analysis of their face tracker on the faces in the Wild database [17]. The database contain images that have varying lighting, resolution, image noise and partial occulusion. Results evaluates that Saragih's Face Tracker exhibits a degree of robustness towards variations encountered in realistic facial video capture [14]. The problem of motion artifacts is resolved to a considerable extent by using this face tracker algorithm, which allowed participants to move their face freely to a certain degree while they were being recorded. Fig. 2 shows an example frame using Saragih's face tracker algorithm.

We have chosen the cheeks from the faces as our region of interest (ROI) for the measurement of heart rate because it tends to be a highly sensitive area of the face and is least likely

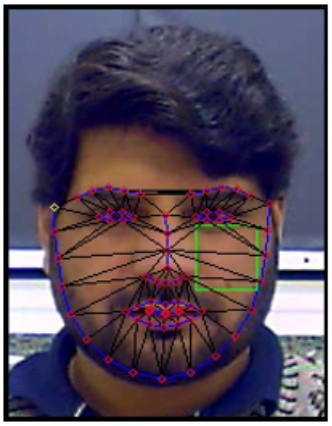


Fig. 2. Saragih's Face Tracker (in blue/black) and ROI extraction (in green).

to be affected by a beard, hair, head covering and glasses. It is also less affected by facial feature movements and is considered the most stable area of the face to extract maximum skin region. We extract a 40 x 40 pixel ROI from the face and one of the points of the Saragih's tracker was used to position it, (see Fig. 2).

C. Spatial Averaging using a Robust Mean Algorithm

We are interested in the RGB color intensities of the skin pixels in the extracted ROI in each frame of the recorded video and therefore we have calculated the average value of each color channel for all pixels in the ROI of each frame. Since the ROI may be affected by illumination changes, shading, non-skin pixels in the ROI and exposure adjustment of the webcam, therefore long-term changes can be found in the averaged ROI values over the period of video capture. In order to compensate for these changes, we have proposed a robust mean algorithm that tries to take into account only the skin color pixels and ignores any non-skin pixels that are affected by lighting and other noisy sources and calculates the average of each color channel for only unaffected skin color pixel values in the ROI of each frame.

Our robust mean algorithm is based on the Mean Shift algorithm [12]. The intuition behind this algorithm is that we will create a small window within the RGB color space and will move that window to the area of maximum pixel density (or a maximum number of pixels). Let us consider Fig. 3, which presents the distribution of pixel colors from the ROI (shown here in 2D for ease of illustration, but calculated in 3D RGB space in the actual algorithm). If we are given a circle *A* indicated on the left with center *A_o* and centroid *A_r*. We have to move the center of the circle to the position of its centroid, which contains maximum pixel distribution. We will keep on moving this circle until the center of the circle matches its centroid and that area will be considered as the area of interest with the maximum number of skin pixels (represented by circle *B*). The algorithm is initialised using the mean of all the points *A_o*. All points within a radius *r* of *A_o* are then used to re-estimate the mean. This is repeated until the mean stabilizes at location *B_o*. The aim is that pixels affected by e.g. specular reflections or the appearance of hair or wrinkles are hence eliminated and only skin color pixels are taken in to account for further processing.

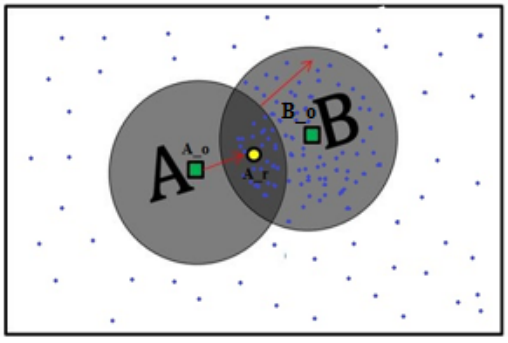


Fig. 3. Robust mean algorithm mechanism: original mean lies at *A_o* and is moved to *B_o* ,which is consider to be the maximum skin pixels area.

We have calculated the mean of each color channel values of all ROI pixels in each frame using our robust mean algorithm. Let us consider $x_j(i)$ as a mean value of color *j*, where *j* represents red, green and blue colors, for the i^{th} frame. The estimated skin color from each frame of the entire video yields three raw signals $x_r(t), x_g(t)$ and $x_b(t)$, representing red, green and blue color traces simultaneously, where *t* denotes time.

D. Detrending

The three RGB signals that we have obtained may still contain some uncertain variations in their average values despite the effect of lighting having been considerably reduced by using the robust mean algorithm. We therefore need to normalize the values of RGB signals in order to achieve further robustness and a cleaner signal. Detrending can eliminate certain variations and normalize the signals for further processing. We have performed this by dividing each sample of these signals by its running average over an interval of at least 20 samples so that it can contain at least one beat period.

$$x'_j(t) = \frac{x_j(i)}{\mu}$$

where *i* represent the i^{th} frame of the signal $j \in \{r, g, b\}$ and μ is the running average over a certain period of time *t*. The three detrended signals consequently obtained are $x'_r(t)$, $x'_g(t)$ and $x'_b(t)$. Fig. 4 shows the effect of detrending on the three RGB signals.

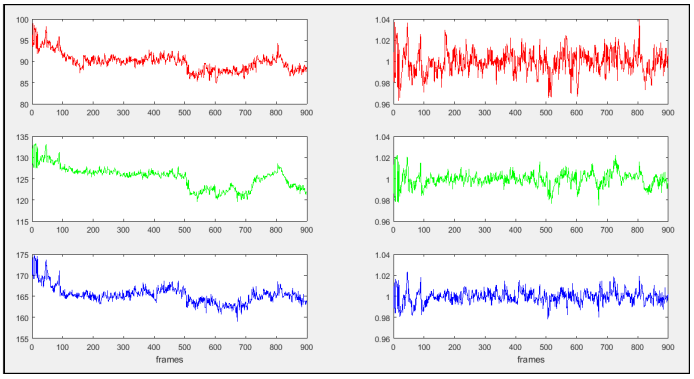


Fig. 4. RGB signals before (left) and after detrending (right).

E. Independent Component Analysis (ICA)

The normalized RGB signals that we have obtained are a mixture of the desired pulse signal and other noise sources. In order to extract the pulse signal from these signals, previous authors have employed ICA at this point [6], which is capable of separating different signals mixed together in different combinations. In our study, we have used joint approximate diagonalization of eigenmatrices (JADE) algorithm of ICA [13]. Conventionally in ICA, the number of outputs is set equal to the number of inputs assigned to it, unless there is a reason to assume otherwise. We have supplied our three normalized RGB signals as an input to the JADE algorithm and as an

output, we obtained three source components, in which one of them will be considered as the pulse signal and the other two are noise sources. We named these three output components as Component 1, Component 2, and Component 3 represented in Fig. 5.

F. PPG Signal Identification

ICA does not reveal any information relating identification of the expected pulse signal among all of three extracted source components. We have employed Poh et al.'s criteria, who have used Fast Fourier Transform (FFT) to obtain the power spectrum of all the extracted components and the component with the highest peak within the operational frequency was selected as the pulse signal [6]. We set the operational frequency range to [0.8-4] Hz, which corresponds to [48-240] bpm, to cover a wide range of acceptable heart rate measurements. The heart rate is then obtained using the following formula: *Heart Rate = Pulse Component highest peak Frequency* x 60.

G. Filter-based Pulse Signal Extraction

Detection of the highest peak within a fixed window is not very reliable as the estimated pulse component may have been affected by noise, resulting in selection of the wrong peak. In order to make it possible for a non-contact heart rate measurement system to work accurately in realistic and more challenging conditions, where lighting and uncertain movements may have been expected, we need a more robust way to extract the heart rate from an estimated pulse component accurately. We have proposed a new method based on the MOSSE tracking algorithm [15], in which we have created a filter using the ground truth pulse signal and the color signal from training videos to improve the extraction of the pulse signal. Let us consider the following equation:

Pulse Signal (oximeter) = Filtered Pulse Component (image)

$$s = f * c \tag{1}$$

where s is the ground truth pulse signal, c is the pulse component obtained from the video images, f is the optimal filter function and * is the convolution operator. For a single pair s and c it is possible to find a function f that satisfies the above equation by converting into the Fourier domain where convolution becomes multiplication i.e.

Equation 1 can be rewritten as:

$$S = FC \tag{2}$$

Where F, S and C are the Fourier transforms of f, s and c respectively. Since we have the pulse signals and pulse components, we can create a filter using the following equation:

$$F = \frac{S}{C} \tag{3}$$

The above filter is only suitable for the specific input signal used to construct it. In the more general case, we can design a filter that can be applied to any of the training signals in order to approximate the desired output signals as closely as possible. We wish to find the filter F which minimizes the error function:

$$\chi^2 = \sum_{i=0}^{N} |F.C_i - S_i|^2 \tag{4}$$

The least squares solution is given by:

$$F = \frac{\sum_i S_i.C_i}{\sum_i C_i.C_i} \tag{5}$$

where S_i is the i^{th} training pulse signal and C_i is the i^{th} training pulse component signal.

We have trained our system using the expected pulse components obtained from the videos and the ground truth pulse signals obtained from oximeter as an input. Using equation (5), we have created a filter and by applying it to each expected pulse component, we obtained an estimated pulse signal. We have employed a leave one out cross validation strategy so that the result of untrained components can be tested against the filter. The average correlation between the filter calculated using all the training data and the filter in each trial is 98 %, which shows that it is not significantly effected by leaving one subject out. The Fourier transform of the estimated pulse signal from input signal C_j is then estimated as follows:

$$K_j = FC_j \tag{6}$$

The frequency corresponding to the highest peak of the Fourier transform of the estimated pulse signal estimates the heart rate and the inverse Fourier transform of the same signal gives us the estimated pulse signal in the time domain.

III. RESULTS

We have evaluated the robustness of our proposed method (described in Section II) to measure the heart rate in the presence of challenging illumination interference and motion artifacts. We have compared the estimated heart rate obtained from recorded video to the ground truth heart rate measured by a pulse oximeter by using the following statistics: mean error (e), absolute mean error ($\bar{e}$), standard deviation of error (σ), root mean square error (RMSE) and correlation coefficient (r). To illustrate the effectiveness and accuracy of our proposed method, we have also compared it with recent state of the art methods for the measurement of heart rate from human face videos.

Fig. 2 presents a sample frame for the tracking of a face from the recorded video and extraction of ROI (cheek) can be shown as a green rectangle. The ROI yields three RGB traces using the robust mean algorithm (Fig. 3, which ignores the pixels affected by specular illumination and taking into account only facial skin color pixels) and calculated the mean of RGB color intensities for all pixels in each frame of the recorded video. The affect of detrending can be viewed in Fig. 4, which normalizes the signals for post processing. The detrended RGB signals were exposed to ICA, resulting in three independent source components represented in Fig. 5. The ICA returns the output source components in random order, it is

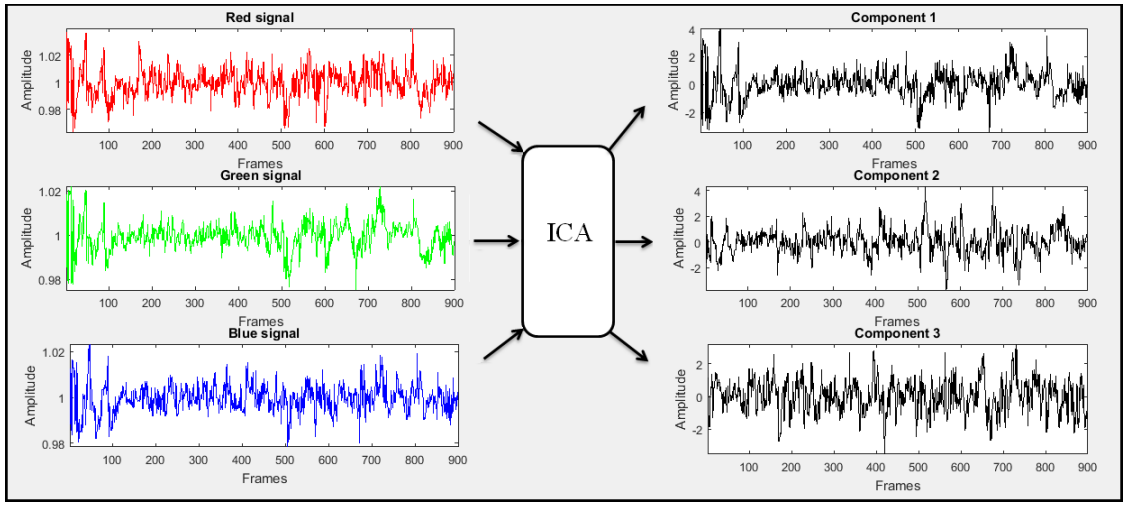


Fig. 5. ICA JADE algorithm decomposes three RGB signals (left) in to three Independent Source Components (right).

therefore not known that which component need to be selected for the estimation of heart rate. The three source components were decomposed from its spatial appearance to its frequency domain using FFT. Fig. 6 clearly shows that component 3 has the highest peak among the three components within the set frequency range, therefore this was chosen as the desired pulse component. The frequency that corresponds to the highest peak of component 3 in its power spectrum within the set frequency range was selected as the heart rate frequency. This was one subject's example for which component 3 was selected as an estimated pulse component, other subjects may have different components which may qualify for the highest peak pulse component.

We can notice in Fig. 6 that we have obtained a prominent peak in the third component but the signal (Fourier domain) is not clean as some noise is still present in it. Our proposed optimal filter method has not only improved the accuracy of the heart rate measured but also resulted in a cleaner pulse signal, which is closer to the measured ground truth pulse signal. Fig. 7 shows the ground truth pulse signal, the chosen estimated heart rate component and the designed filter. The filter, when applied to the expected heart rate component revealed the estimated pulse signal as presented in Fig. 7. We can notice that the estimated filtered pulse signal is more periodic in nature as compared to the estimated pulse component. The noise is hence filtered and we get a cleaner periodic signal, which is a more true representation of a pulse signal.

Fig. 8 compares the FFT power spectrum of pulse signals obtained before and after the application of the optimal filter. The power spectrum of both of these signals shows that the heart rate peak (indicated by a red arrow) become more prominent after we applied the filter to it. The signal noise is also reduced, which shows that the affect of illumination interference and motion distractions is reduced and we get a cleaner pulse signal.

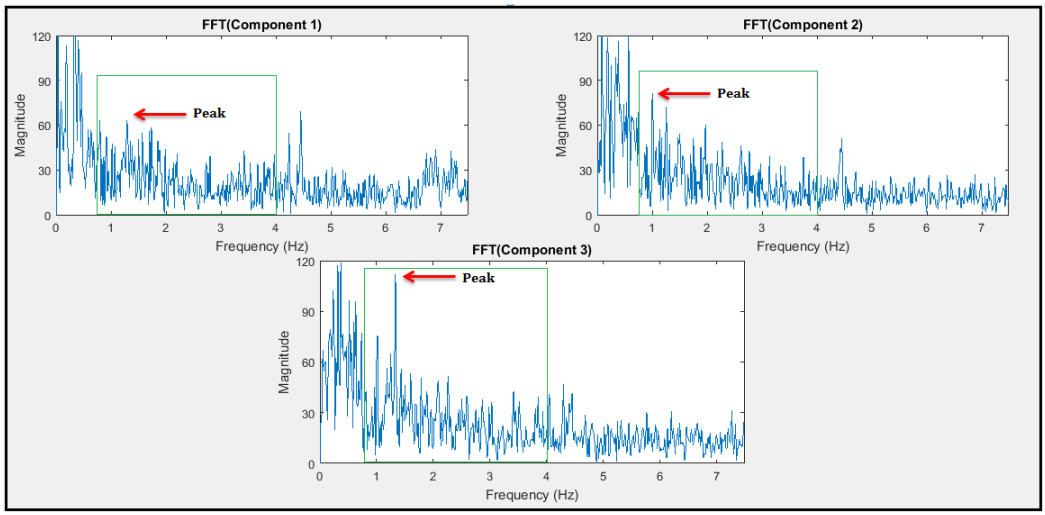


Fig. 6. FFT of Component 1 (top left), Component 2 (top right), and Component 3 (Bottom one with the highest peak within set frequency range) . Green rectangle represents the set frequency range. Red arrow indicates heart rate peak.

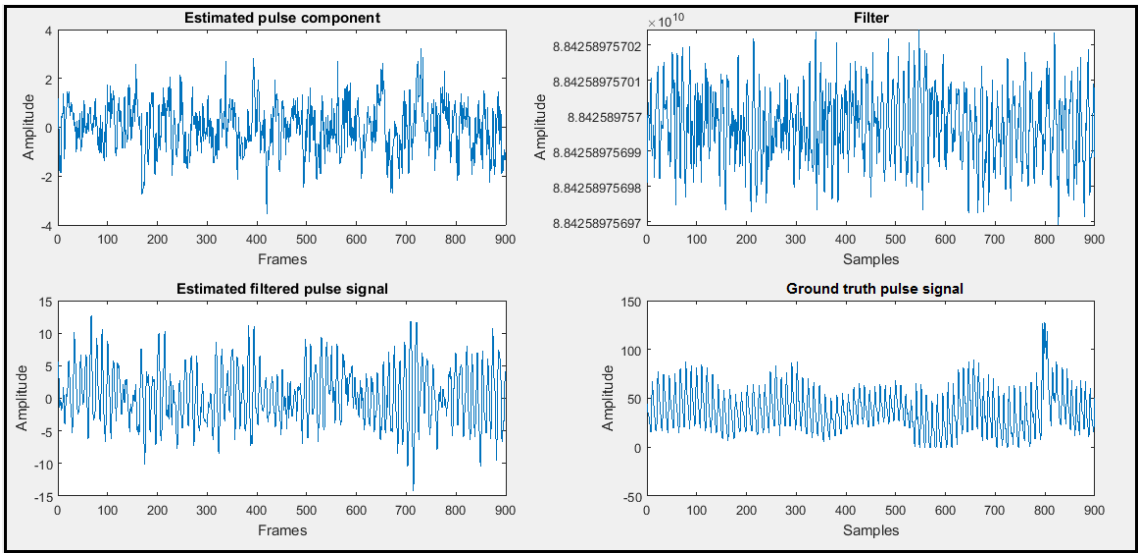


Fig. 7. Filter (top right) applied to Estimated pulse component (top left), we obtained Estimated filtered pulse signal (bottom left),which look clean and more periodic like Ground truth pulse signal (bottom right).

Table I shows the effectiveness of our proposed robust mean algorithm, which we have used instead of using the mean of all pixels. We can notice that robust mean algorithm reduces the absolute mean error and therefore has a positive affect on improving the overall results. We can therefore assume that the robust mean algorithm ignores pixels effected by illumination and calculating the mean of only those pixels that are mostly skin pixels. Fig. 9 presents the comparison of different processing steps and their impact on the improvement of the measured heart rate. We can notice that the green line, which is the final step yielding the estimated heart rate from facial videos using filtered robust mean method, come closer to the ground truth heart rate.

Table II presents the robustness of our proposed method compared to other state of the art methods. It describes statistical evaluation of our proposed method and some recent state of the art methods. We can notice that our method has the minimum mean error, absolute mean error, standard deviation of error, and RMSE as compared to the results of other methods, therefore, it shows that our proposed method performed better under more realistic and challenging conditions. Fig. 10 illustrates the correlation coefficient plots of our proposed method in comparison to other methods. The higher correlation coefficient of 0.82 of our proposed method demonstrates that the heart rate measured from videos is highly correlated to the ground truth heart rate as compared to other methods. This confirms that our proposed method has an accuracy of 82 % and is the most robust method for the measurement of heart rate and extraction of a clean pulse signal under more realistic conditions.

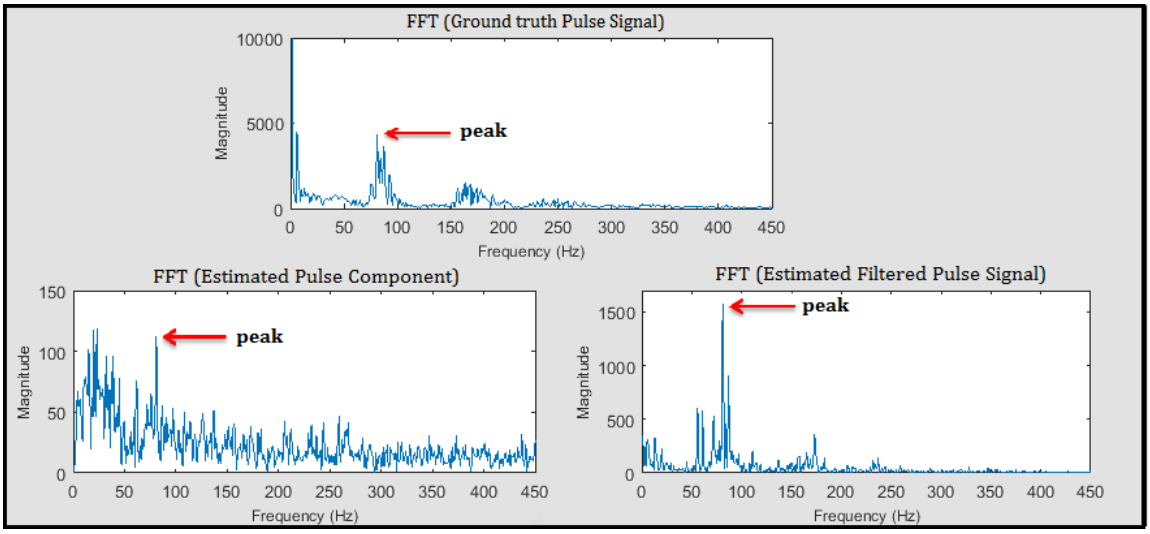


Fig. 8. FFT of Ground truth pulse signal (top), FFT of estimated Component(bottom left) and FFT of Estimated filtered pulse signal (bottom right), red arrow indicates heart rate peak.

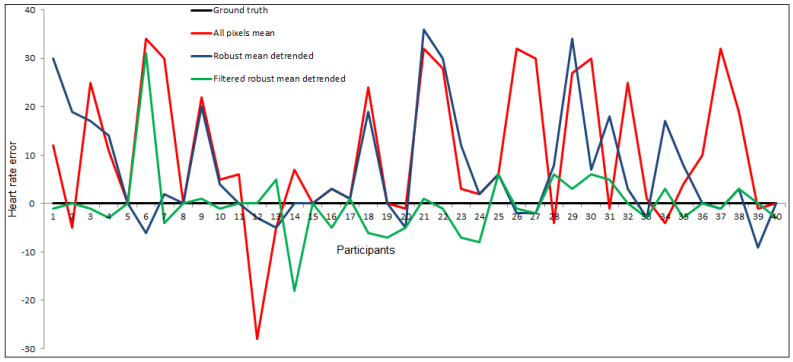


Fig. 9. Plot of comparison of different processing steps measured heart rate against the ground truth for 40 participants. Black line is a ground truth heart rate, red line indicated the use of all pixels mean, blue line shows the impact of robust mean algorithm. Green line represents our proposed method and shows that it remain closer to the black line of ground truth.

IV. DISCUSSION

Measurement of the heart rate from human face videos captured by means of a webcam under realistic scenarios is a challenging task as we have found that most of the previous methods are dependent on highly controlled conditions. In order to employ this technology in real world applications, it needs to be robust and easily adaptable to a wide range of capturing scenarios. We have addressed these problems in our work and tried to overcome them. We have therefore made our experimental environment challenging, with difficult lighting and some natural facial motion. Our results demonstrates the improved performance of our proposed method but there are some concerns that need to be taken in to account and require further improvement. The ROI was the cheeks in our system but we could also investigate other parts of the body that could give us even better estimation of variations in skin color pixels with respect to each heartbeat. We have normalized our RGB signals before presenting them to ICA, but if we can further do some pre-processing normalization of our raw signals then it may perform better in the post ICA stage, where we are identifying expected pulse signal. We have chosen the independent component with the highest peak in the power spectrum within the set frequency range, but we have found in some components that a smaller peak was more prominent (high compared to its neighbours) than the highest peak components and the heart rate obtained was closer to the ground truth heart rate. The criterion for selecting the correct source component for estimating heart rate may need to be improved by setting some threshold. We found that our method performed less efficiently in the cases of very dark skin color subject and where the participant wear heavy make up. The expected solution to this problem can be the use of near-infrared cameras, so that if the heart rate can not be measured by the skin pixels color variation of a webcam then thermal energy radiated from skin using near-infrared cameras can cover it. The ground truth data which we get from the pulse oximeter can be made more robust if we can obtain it from using ECG signal which is considered more accurate. This could result in obtaining more accurate heart rate measurements and a cleaner pulse signal.

TABLE I
RESULTS OBTAINED FROM OUR PROPOSED ROBUST MEAN ALGORITHM VS ALL PIXELS MEAN ALGORITHM

	All pixels mean algorithm	Robust mean algorithm
e	-2.17	6.92
$\bar{e}$	11.22	8.72
σ	14.28	10.16
ϵ	18.02	13.29
r	0.14	0.53

The table shows mean error e, absolute mean error $\bar{e}$, standard deviation of error σ, root mean square error ϵ , and pearson correlation r.

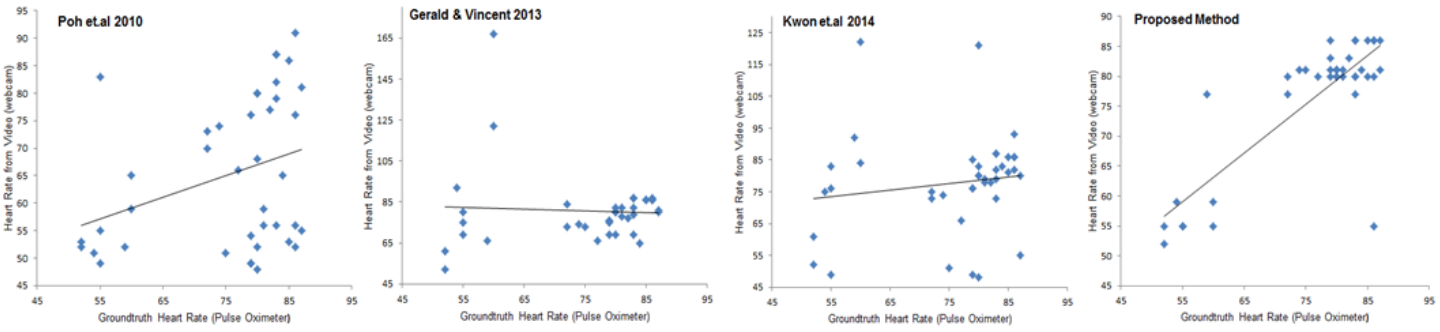


Fig. 10. Correlation plots for our proposed method in comparison to other methods for a total of 40 participants. Our proposed method plot shows that the estimated heart rate from videos is highly correlated to the ground truth heart rate.

TABLE II
RESULTS OBTAINED FROM OUR PROPOSED METHOD IN COMPARISON WITH OTHER HEART RATE MEASUREMENT METHODS

	Poh *et al*. 2010	Gerald 2013	Kwon *et al*. 2012	**Proposed**
e	10.30	-5.15	-2.17	**-0.22**
$\bar{e}$	12.75	10.30	11.22	**3.77**
σ	12.43	19.67	14.28	**5.57**
ϵ	17.70	21.99	18.02	**6.67**
r	0.33	-0.05	0.14	**0.82**

The table shows mean error e, absolute mean error $\bar{e}$, standard deviation of error σ, root mean square error ϵ , and pearson correlation r.

V. CONCLUSION

We have presented an improved method for the measurement of heart rate from human face videos under realistic and more challenging conditions. Evaluation shows that our proposed method not only improved on the existing methods for the measurement of the heart rate but also resulted in a better pulse signal, which could be integrated in to many other real life applications such as human biometrics, emotion recognition, driving monitoring systems, and medical care equipment.

REFERENCES

[1] Scanlon, Valerie C. and Tina Sanders. *Essentials of anatomy and physiology.* F.A. Davis Company, 2007.

[2] I. Pavlidis, J. Dowdall, N. Sun, C. Puri, J. Fei, and M. Garbey "Interacting with human physiology," *Computer Vision and Image Understanding*, vol. 108, 2007, pp. 150-170.

[3] M. Garbey, N. Sun, A. Merla, and I. Pavlidis,"Contact-free measurement of cardiac pulse based on the analysis of thermal Imagery," *IEEE Transactions on Biomedical Engineering*, vol. 54, 2007, pp. 1418-1426.

[4] R.G. Gatto, "Estimation of instantaneous heart rate using video infrared thermography and ARMA models," Ph.D. thesis, University of IIIinois at Chicago, 2009.

[5] C. Takano and Y. Ohta, "Heart rate measurement based on a time-lapse image,"*Medical Engineering and Physics*, vol. 29, no. 8, 2007, pp. 853-857.

[6] 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, 2010, pp. 10762-10774.

[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," *ACM Transactions on Graphics (TOG)*, vol. 31, no. 4, 2012, pp. 1-8.

[8] G. de Haan and V. Jeanne, "Robust pulse-rate from chrominace-based rPPG," *IEEE Transactions on Biomedical Engineering*, vol. 60, 2013, pp. 2878-2886.

[9] J. Allen, "Photoplenthysmography and its application in clinical physiological measurement," *Physiological Measurement*, vol. 28, 2007, pp. 1-39.

[10] P. Viola and M. Jones, "Robust real-time face detection", *Computer Vision*, vol. 57, no. 2, 2004, pp.137-154.

[11] J.Saragih (2015), Face tracker. Available: https://github.com/kylemcdobald/FaceTracker. Last accessed 10/03/2015.

[12] Z. Wen and Z. Cai, "Mean shift algorithm and its application in tracking of objects", *2006 International Conference on Machine Learning and Cybernetics*, pp. 4024-4028.

[13] J. Cardoso, "High-order contrasts for independent component analysis (ICA)", *Neural Computation*, vol. 11, 1999, pp. 157-192.

[14] J. Saragih, S. Lucey, and J. Cohn, "Face alignment through subspace constrained mean-shifts", *IEEE International Conference on Computer Vision (ICCV), 2009.*

[15] D S. Bolme, J.R. Beveridge, B.A. Draper, and Y.M. Lui, "Visual object tracking using adaptive correlation filters." *Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2544-2550.

[16] S. Kwon, H. Kim, and K.S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone", *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 2174-2177.

[17] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments", *Technical Report 07-49, University of Massachusetts, Amherst*, 2007.

Muhammad Waqar received the B.Sc. and M.Sc. degrees from University of Peshawar, Pakistan. He did M.S in Telecommunication Engineering from University of Sunderland, UK. He is currently doing Ph.D. degree in computer science from Aberystwyth University, Wales, UK. His research interests include image analysis, motion estimation algorithms,feature-detection techniques and computer vision.

Reyer Zwiggelaar received the Ir. degree in applied physics from the State University Groningen, Groningen, The Netherlands, in 1989, and the Ph.D. degree in electronic and electrical engineering from University College London, London, U.K., in 1993. He is currently a Professor at the Department of Computer Science, Aberystwyth University, U.K. He is the author or co-author of more than 150 conference and journal papers. He is an Associate Editor of the Journal of Biomedical and Health Informatics (formerly IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE). His current research interests include medical image understanding, especially focusing on mammographic and prostate data, pattern recognition, statistical methods, texture-based segmentation, and feature-detection techniques.

Bernard Tiddeman is a Reader and Head of Department in Computer Science at Aberystwyth University. He obtained his BSc from University of St Andrews in 1992, MSc from Manchester University in 1994 and PhD from Heriot-Watt University in 1998. From 1999-2010 he worked as a researcher and then lecturer at the University of St Andrews. His research interests include 2D and 3D facial image analysis and synthesis, texture modelling for age estimation and age progression, skin health analysis and medical image analysis.