# Recognition of Correct Finger Placement for Photoplethysmographic Imaging

Walter Karlen[1,2], Joanne Lim[1], J. Mark Ansermino[1], Guy A. Dumont[1], and Cornie Scheffer[2]

*Abstract*— In mobile health applications, non-expert users often perform the required medical measurements without supervision. Therefore, it is important that the mobile device guides them through the correct measurement process and automatically detects potential errors that could impact the readings. Camera oximetry provides a non-invasive measurement of heart rate and blood oxygen saturation using the camera of a mobile phone. We describe a novel method to automatically detect the correct finger placement on the camera lens for camera oximetry. Incorrect placement can cause optical shunt and if ignored, lead to low quality oximetry readings. The presented algorithm uses the spectral properties of the pixels to discriminate between correct and incorrect placements. Experimental results demonstrate high mean accuracy (99.06%), sensitivity (98.06%) and specificity (99.30%) with low variability. By sub-sampling pixels, the computational cost of classifying a frame has been reduced by more than three orders of magnitude. The algorithm has been integrated in a newly developed application called *OxiCam* where it provides real-time user feedback.

## I. INTRODUCTION

Pulse oximetry provides a non-invasive measurement of heart rate (HR) and blood oxygen saturation ($SpO_2$). Blood perfused tissues are illuminated by a light source with a known spectrum, and a photo detector records the photons that were not absorbed or scattered and passed through the tissue. Beer-Lambert's law stipulates that absorption of a photon in a given medium $i$ is based on the wavelength of the photon $\lambda$, medium concentration $c$ and distance $d$ traveled, such that

$$I = I_0 e^{-\sum_{i=1}^{n} \epsilon_i(\lambda)c_i d_i}, \tag{1}$$

where $I$ is the light intensity and $\epsilon$ the extinction coefficient. Since the optical path of the photons and consequently the total absorption changes with each heart beat due to an increase in blood volume, the light intensity measured by the sensor also changes. The recorded variation of light intensity is the photoplethysmogram (PPG) which allows for accurate estimation of the heart rate (Fig. 1). When the photo detector used in pulse oximetry is replaced by a pixel array, it is called photoplethysmographic imaging [1], and when $SpO_2$ is also calculated, it is called camera oximetry [2], [3].

We have developed an Android mobile phone software application called *OxiCam* that records HR and $SpO_2$ using
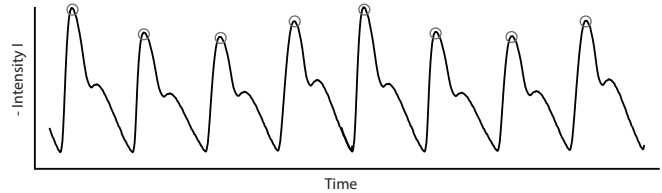
Fig. 1. Photoplethysmogram obtained from pulse oximetry. When blood volume in the finger is maximal, the recorded light intensity is lowest due to increased absorption (circles).
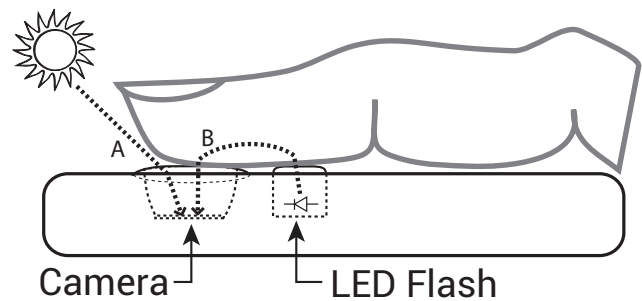


Fig. 2. Typical setup for photoplethysmographic imaging. A finger is placed on the mobile phone camera and continuously illuminated using the integrated LED flash. Photons not passing through oxygenated tissue or coming from other light sources create an optical shunt (A). Ideally, all recorded photons have passed oxygenated tissue before hitting the sensor (B).

camera oximetry. The intended uses of *OxiCam* are mobile health applications for non-specialized health workers and lay users interested in their own health. Since the target users are not experts in pulse oximetry or clinical monitoring in general, it is important that the device or the software can guide them through the correct measurement process and automatically detect potential errors that could impact the readings.

In this paper, we describe a novel method employed by *OxiCam* to automatically detect correct finger placement on the camera lens and provide feedback to the user. Incorrect placement may cause optical shunt and if ignored, lead to low quality oximetry readings [4]. Optical shunt occurs when photons that do not pass through tissue containing oxygenated blood are detected by the sensor (Fig. 2). This creates an undesirable, low signal-to-noise ratio and depending on severity, introduces errors in the HR and $SpO_2$ calculations. We present decision criteria for automatically distinguishing correct from incorrect finger placement and conduct an experiment to optimize the decision thresholds.

## A. Background

The current literature describes methods based on the spectral properties of the image to discriminate between correct or incorrect placement of the finger. In [5] a relatively simple method was presented. Seven rules, based on the mean $\bar{x}$ and standard deviation $\sigma$ of the three color channels $r$, $g$ and $b$, were used together with 5 thresholds $Th$ such as

$$p(\bar{x}) = \begin{cases} 1 & \begin{aligned} &\text{if } \bar{x}_g + \sigma_g \geq Th_g^{min} \\ &\text{and } \bar{x}_r - \sigma_r > Th_r^{min} \\ &\text{and } \bar{x}_g + \sigma_g < Th_g^{max} \\ &\text{and } \bar{x}_b + \sigma_b < Th_b^{max} \\ &\text{and } \sigma_r \geq Th_\sigma^{min} \\ &\text{and } \sigma_g \geq Th_\sigma^{min} \\ &\text{and } \sigma_b \geq Th_\sigma^{min} \end{aligned} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where $p$ is the label for correct (=1) and incorrect (=0) finger placement, $\bar{x}$ is a vector containing 8-bit pixel values of length $N$. The thresholds were determined during a calibration phase of undisclosed duration at the beginning of each recording. Quantitative evaluation of the correct finger placement detection rate was not reported. The reported method is not computationally efficient, as the means and standard deviations of all pixels have to be computed at each new image frame.

The goal of the subsequent sections is to present an improved detection method and to quantify the finger recognition rate.

## II. METHODS

### A. Data Collection

Twenty video recordings from a total of 5 subjects, each of duration of 30 s, were collected with the *OxiCam* application. *OxiCam* was installed on a Samsung Galaxy Ace mobile phone, running the Android 2.3.4 operating system. The videos were saved in 240 x 320 pixels resolution (QVGA) and the "mp4" format to minimize processing load. The video sampling rate was 30 Hz. The white balance was set to "incandescent" as this has been shown to be the optimal configuration for this type of camera [6]. Each recording consisted of an initial phase of 15 s where the subject's finger was placed incorrectly, followed by 15 s of correct placement on the camera lens and LED. Incorrect placement was defined as a placement that would produce optical shunt. Various background colors and brightness levels were simulated during the first phase. Incorrect finger placement scenarios were produced by systematically changing the optical shunt from 20% to 100%. Annotations for correct finger placement were verified manually by the first author.

### B. Algorithm Design

It was clear from the recorded videos that the correctly placed finger presents a unique spectral distribution which is dominant in the red color spectrum that was also observed by Lamonaca et. al. [5]. Consequently, the histogram for each

frame was calculated and displayed over time (Fig. 3). It was obvious that correct finger placement could be distinguished from incorrect placement using 3 thresholds. Blue and green were never present in high concentrations while red alone was always present in high concentrations. The three new classification rules established were:

$$p(\bar{x}) = \begin{cases} 1 & \begin{aligned} &\text{if } \bar{x}_r > Th_r \\ &\text{and } \bar{x}_g < Th_g \\ &\text{and } \bar{x}_b < Th_b \end{aligned} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The 3 thresholds are determined by :

$$Th_\lambda = \bar{y}_\lambda + f \times \sigma_\lambda \quad (4)$$

where $\bar{y}$ is a frame vector of known videos of correct placements and $f$ is a variable scaling factor. This rule was inspired by (2) that was established by [5]. We have introduced the scaling factor $f$ as it allows for adjustment of the confidence intervals for determining the mean intensity of each color channel in a new, unknown frame.

### C. Parameter Optimization

To find the optimal thresholds, we performed a 'leave-one-out' cross validation experiment. Videos were divided into a training set and a test set. The training set was composed of all recordings but one. The remaining video was used in the test set. The test video was then exchanged with a video in the training set. This was repeated until each video had been in the test set once. During training, the thresholds were modified by changing $f$ from 0.01 to 2 in increments of 0.01. This corresponds to a confidence interval of 0.8% to 95%. The threshold providing the highest performance on the training set was selected for performance testing on the test set. The performance function $P$ was designed such that

$$P = Se + 2 \times Sp \quad (5)$$

$$Se = \frac{TP}{TP + FN} \quad (6)$$

$$Sp = \frac{TN}{TN + FP} \quad (7)$$

where $Se$ is the sensitivity, $Sp$ the specificity and $TP$ is the number of true positives, $TN$ is the number of true negatives, $FN$ the number of false negatives, and $FP$ the number of false positives. Specificity was weighted more as it is more important in everyday use to have fewer $FP$ (type 1 errors). It is preferable to ignore frames that have correct placement compared to calculating HR and SpO$_2$ during an incorrect placement. Given the low degree of freedom of the search space, over-fitting during training was not likely to occur and no additional validation on the training procedure was necessary.

Using the complete set of pixels for each frame to calculate the finger placement is computationally expensive. We were interested in reducing this computational effort. For this, we repeated the experiments by using only a fraction of the
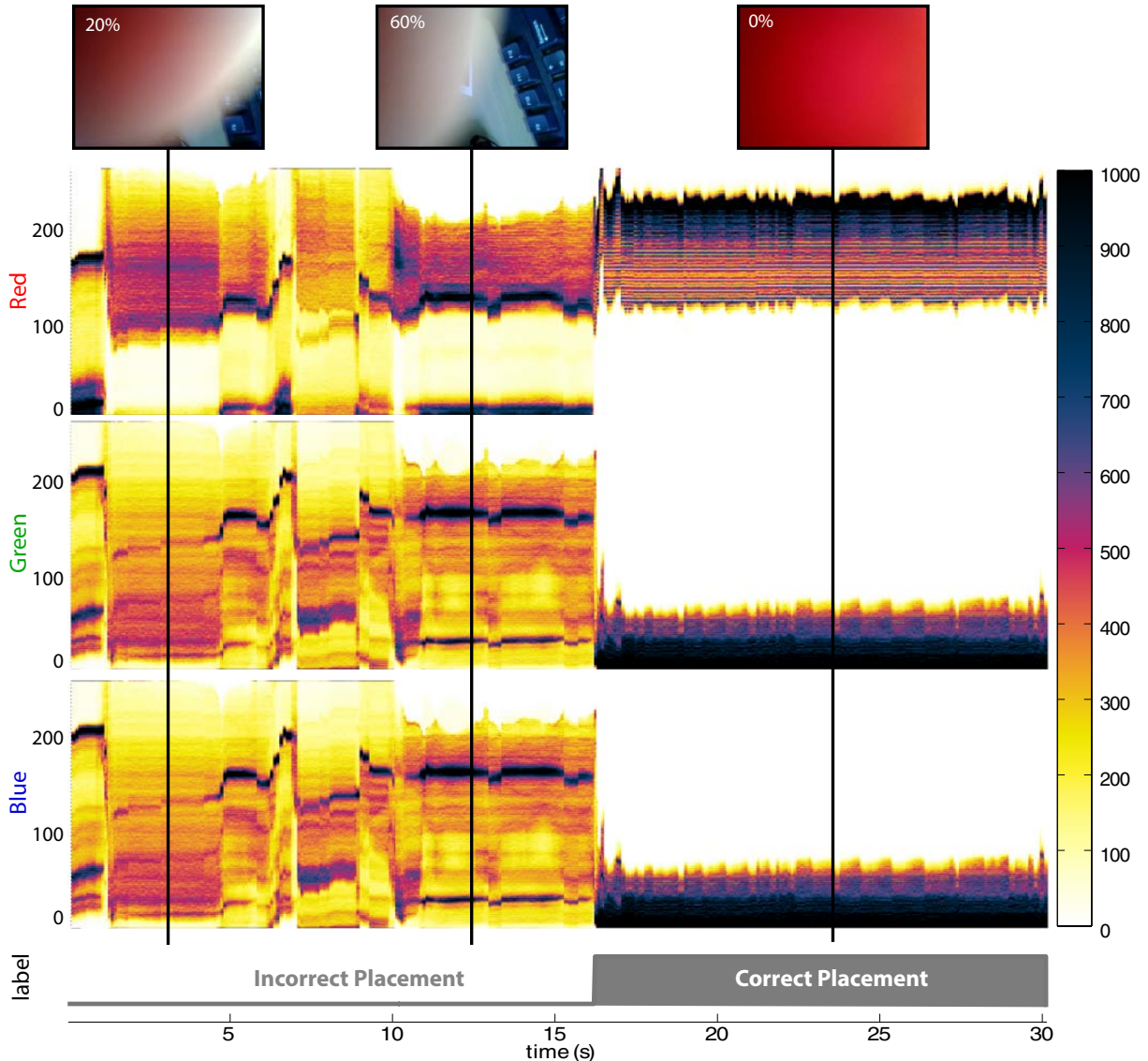
Fig. 3. Histogram of the red (top), green (middle), and blue (bottom) video channel for a typical video recording. The boxes show three video frames at a given time. In the upper left corner of these boxes the estimated optical shunt is displayed.

pixels in the test set. At each time step, pixels were randomly selected for the calculation of the mean $\bar{x}$. The number of pixels $N$ was modified to 10%, 1%, 0.1%, and 0.01% of the original number of pixels in a frame.

We report the mean and standard deviation of the mean error as well as the sensitivity and specificity. Sensitivity and specificity were also represented in a receiver operating curve (ROC) graph.

## III. RESULTS

The median scaling factor that produced the best training performance was $\tilde{f} = 0.7$. This $\tilde{f}$ was selected to calculate the generalized thresholds. As desired, these thresholds produced a higher specificity than sensitivity (Fig. 4). A high classification rate with low variability over the 20 iterations was obtained (Table I).

TABLE I
ACCURACY, SENSITIVITY AND SPECIFICITY OF OPTIMAL CLASSIFIER

| Pixels | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|--------|--------------|-----------------|-----------------|
| 10% | $99.43 \pm 0.12$ | $98.80 \pm 0.37$ | $99.26 \pm 0.76$ |
| 1% | $99.29 \pm 0.16$ | $98.53 \pm 0.42$ | $99.31 \pm 0.74$ |
| 0.1% | $99.06 \pm 0.20$ | $98.17 \pm 0.41$ | $99.30 \pm 0.74$ |
| 0.01% | $98.51 \pm 0.22$ | $97.58 \pm 0.35$ | $99.07 \pm 0.32$ |

### A. Computational Benefit

The rules as described in (2) require $9 \times N + 9$ operations for each frame, where $N = 76800$ is the number of pixels in each frame. The newly proposed rules (3) require $3 \times N + 3$. This decrease by a factor of 3 is not significant but can be helpful when processing power should be conserved. The selection of a subsample of pixels to $10^3$ of the original frame size can save a significant number of operations without
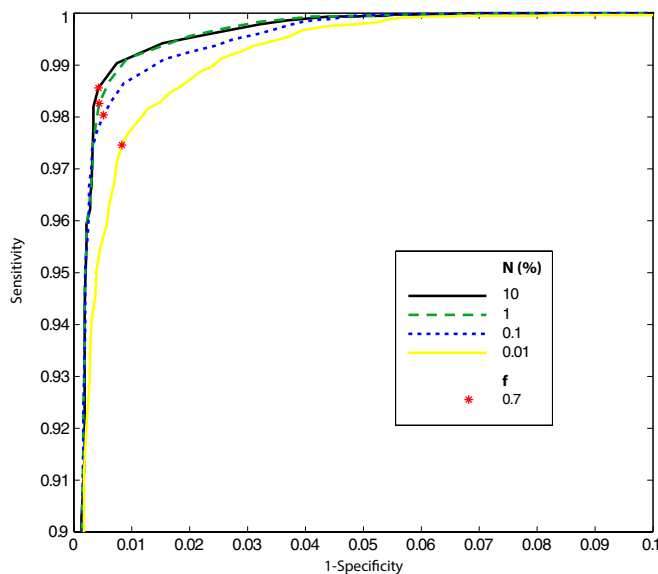
Fig. 4.  Receiver operating curve (ROC) for 4 different pixel vector lengths $N$. A perfect classifier would lie in the upper left corner (0,1). The stars indicate the test classifier performance when $f = 0.7$ (best performance $P$ during training). Note that for resolution reasons the axes show only a narrow range of 0.9 to 1 for sensitivity and specificity.

compromising performance (Table I). Further reduction in number of pixels causes a visible reduction in performance (Fig. 4).

## IV. DISCUSSION

We have demonstrated an efficient algorithm to assess the accuracy of finger placement on the camera lens of a mobile phone. This is an essential approach for real-time camera oximetry. High sensitivity and specificity has been achieved. The presented method is a generalized solution and does not require real-time adaptation of parameters. This generalization has been tested on identical cameras and white balance setting only. Changes in white balance will inevitably alter spectral composition of captured images [6]. Under such conditions, the thresholds will require adjustments to be made.

The suggested decision rules reduced the computational effort by a factor of 3 compared to previously published methods [5] by eliminating the real-time computation of the standard deviation of all pixels. Additionally, we have successfully shown that the computational load can also be reduced by three orders of magnitude by sub-sampling the frame without impacting the classification accuracy. This method can be optimized further by pre-selecting pixels from regions-of-interest. Further, the introduction of a time dependent algorithm that uses previous frame classifications could eliminate noise in the classification, particularly in state transitions when the finger is placed onto the lens.

The algorithm functionality could be extended with pulse recognition. The additional processing step would provide not only information about correct placement, but also indicate if correct pressure is applied. This could be achieved by estimating the quality of each pulse [7]. Such an approach
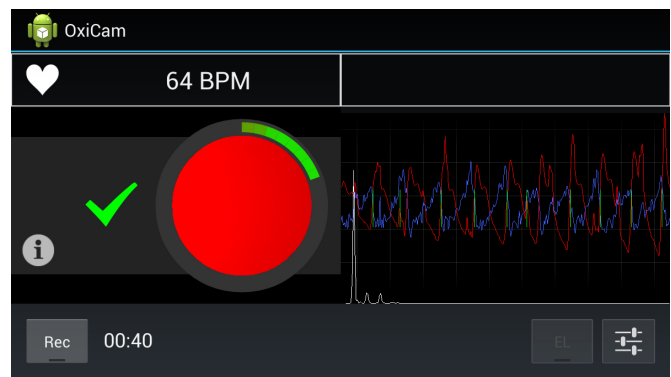


Fig. 5.  *OxiCam* user interface with real-time feedback about correct finger placement (green checkmark).

could guide the user more accurately in getting a good PPG signal.

The presented algorithm has been implemented in *OxiCam*. A green checkmark is provided as feedback for correct placement in real-time, frame-by-frame to the user (Fig. 5). *OxiCam* is currently undergoing clinical testing for HR and SpO$_2$ accuracy.

## REFERENCES

[1] T. Wu, V. Blazek, and H. J. Schmitt, "Photoplethysmography imaging: a new noninvasive and noncontact method for mapping of the dermal perfusion changes," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, P. Brouwer, I. Gannot, T. I. Karu, I. J. Bigio, V. V. Tuchin, A. V. Priezzhev, S. Andersson-Engels, K. Koenig, H. J. Tanke, J. G. Fujimoto, P. A. Oberg, G. J. Mueller, R. Lubart, Y. V. Gulyaev, T. G. Papazoglou, G. J. Puppels, H. Schneckenburger, R. W. Steiner, C. F. P. van Swol, and K. Svanberg, Eds., vol. 4163, Nov. 2000, pp. 62–70.

[2] K. Humphreys, T. Ward, and C. Markham, "A CMOS Camera-Based Pulse Oximetry Imaging System." in *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 4, Jan. 2005, pp. 3494–7.

[3] W. Karlen, J. Lim, J. M. Ansermino, G. A. Dumont, and C. Scheffer, "Design Challenges for Camera Oximetry on a Mobile Phone," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 2448–51.

[4] P. D. Mannheimer, "The light-tissue interaction of pulse oximetry." *Anesthesia and analgesia*, vol. 105, no. 6 Suppl, pp. S10–7, Dec. 2007.

[5] F. Lamonaca, Y. Kurylyak, D. Grimaldi, and V. Spagnuolo, "Reliable pulse rate evaluation by smartphone," *2012 IEEE International Symposium on Medical Measurements and Applications Proceedings*, no. 1, pp. 1–4, May 2012.

[6] W. Karlen, J. M. Ansermino, G. A. Dumont, and C. Scheffer, "Region of Interest for Camera Oximetry," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013, p. accepted.

[7] W. Karlen, K. Kobayashi, J. M. Ansermino, and G. A. Dumont, "Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation." *Physiological measurement*, vol. 33, no. 10, pp. 1617–29, Sept. 2012.