3-15-2013

Constrained independent component analysis for non-obtrusive pulse rate measurements using a webcam

Survi Kyal

Constrained Independent Component Analysis for Non-Obtrusive Pulse Rate Measurements Using a Webcam

by

Survi Kyal

Master of Science Thesis

In

Electrical Engineering

Approved by:

_____

(Gill R Tsouri, Ph. D.)

_____

(Eli Saber, Ph. D.)

_____

(Sohail A. Dianat, Ph. D.)

Electrical and Microelectronic Engineering Department

Kate Gleason College of Engineering

Rochester Institute of Technology

Rochester, New York

sxk8835@rit.edu

Date: March 15th, 2013

Contents

Abstract

Assessment of cardiac function of a patient is very important for understanding a patient's physiological state. Remote measurements of the cardiac pulse can provide comfortable physiological assessment by minimizing the amount of wires and cables and allowing for near continuous measurements. It has been found that state-of-the-art algorithms based on independent component analysis (ICA) suffer from a sorting problem which hinders their performance. This effect is demonstrated in this work. The automated pulse detection techniques are applied to RGB color video recordings of the facial region of a person being monitored for cardiac function in a remote sensing environment. Automated face tracking is employed to locate the region of interest and address motion artefacts.

This work proposed and evaluates a novel algorithm based on constrained source separation, aka, constrained independent source separation (cICA) to accurately estimate the pulse rate of a patient by solving the sorting problem observed in the ICA based approach. The constrained optimization problem incorporates prior information and additional requirements in the form of constraints. A reference signal with a single tone frequency corresponding to a possible heart rate is fed to the cICA algorithm. This forces the output signal to match the reference signal embodying prior knowledge about an underlying IC. It is also shown that with this algorithm a near photoplethysmography (PPG) signal corresponding to the variations in blood volume in the body can be extracted. An IRB approved study encompassing 45 subjects resulted in Bland-Altman analysis with an FDA-approved finger blood volume pulse (BVP) sensor demonstrating that the proposed algorithm provides significantly improved accuracy.

List of Figures

Summary of Contributions

The following is a list of contributions presented in this work.

- **Algorithm Implementation:** The proposed constrained Independent Component Analysis (cICA) algorithm was implemented and optimized in Matlab and was used to analyze 45 video streams captured in a study. State of the art algorithms using the traditional ICA approach were implemented as well to form a basis for comparison via Bland-Altman plots.
- **Publication:** G. R. Tsouri, S. Kyal, S. Dianat, L. K. Mestha, "Constrained-ICA Approach to Non-Obtrusive Pulse Rate Measurements", published in SPIE Journal of Biomedical Optics, Vol. 17(7), and August 2012.
- **Patent:** L. K. Mestha, S. Kyal, G. R. Tsouri, S. Dianat, Xerox Corp., "Estimating Cardiac Pulse Recovery From Multi-Channel Source Data Via Constrained Source Separation", US 13/247,683 Patent Pending, Sept. 2011.

1 Background - ICA

ICA is a widely used technique in blind source separation, feature extraction and signal detection. It transforms a multivariate random signal into a signal having mutually independent components. It assumes that the observed signals $x(t)$ are the instantaneous linear mixture of independent sources$s(t)$, i.e.

$$x(t) = A\, s(t) \tag{1}$$

where, $x(t) = [x_1(t)x_2(t)x_3(t)]^T, s(t) = [s_1(t)s_2(t)s_3(t)]^T$ are vectors and A is a 3x3 square matrix (considering three sources are mixed) containing the mixture coefficients. In this case, the number of sources is equal to the number of observations. ICA aims at extracting the original sources $s(t)$ by estimating the demixing matrix W which is an approximation of the inverse of the original mixing matrix A.

$$y(t) = W\, x(t) \tag{2}$$

$y(t)$ is an estimate of the underlying source signals $s(t)$. The overall ICA process is described as follows:

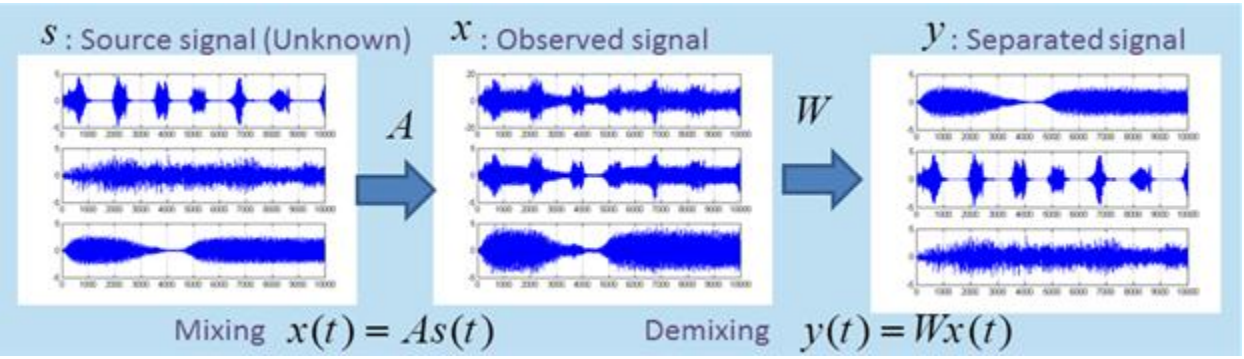


Figure 1: Source separation using Independent Component Analysis

According to central limit theorem, under given certain conditions, distribution of the mixture of independent and identically distributed (i.i.d.) random variables converges to the

Gaussian distribution as number of sources increases (refer figure 2). Estimating independent components corresponds to finding a demixing matrix W such that each entry of $y = Wx$ is as non-gaussian as possible. This implies that to recover independent source signals, W should maximize non-gaussianity of each source. A cost function which measures the non-gaussianity is maximized iteratively. Negentropy is used to separate ICs from their mixtures since the sources considered in source separation usually have non-gaussian distributions.

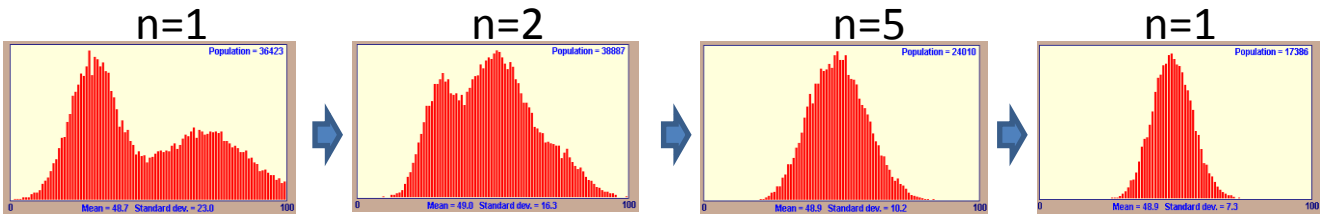


Figure 2: Central Limit Theorem

In the framework of pulse rate measurements, the periodic variation in color due to the pulsating blood flow beneath the skin surface controlled by the heart is the point of interest. Other sources, such as movement of the subject, distance of the camera from the subject or fast variations in ambient light, are mixed with the pulsating heart and acts as interference in the pulse rate measurements.

It has been observed that the standard ICA technique suffers from the sorting problem. That means the interested component could be present in any of the ICA output channels. This sorting problem was recognized in [2, 3] and solved either by selecting the second component always [2] or by selecting the peak frequency of the IC with the highest power [3]. Moreover, in the latter case, the selected component would vary for different measurements.

2 Previous Work

2.1 Heart rate measurement based on a time-lapsed image [1]

In this work, the pulse rate was measured using a time-lapsed image acquired from a CCD camera. A part of the subject's skin (cheek) chosen as the region of interest (ROI) was constantly captured for 30 sec. The changes in the average image brightness of the ROI were measured using image-processing software and processed by a series of operations. As sampling of the time-lapsed image was not conducted with a constant time interval, the measured illumination data were interpolated and regulated by up-sampling software. Different interpolation techniques including the cubic spline, the B-spline and the non-uniform rational B-spline were tested and the first two performed well. This was followed by operations involving first-order derivative, a low pass filter of 2 Hz and a sixth-order auto-regression spectral analysis on the interpolated data. Other spectral analysis like the fast Fourier transform and the wavelet transform were evaluated too. Using the wavelet transform and AR spectral analysis, a clear peak could be seen at appropriate heart rate frequency. Fourteen subjects participated in the experiment and a correlation coefficient of 0.90 was obtained for the measurement of heart rate.

2.2 Non-contact, automated cardiac pulse measurements using video imaging and blind source separation [2]

In this work, a new methodology was introduced which could be applied to color video recordings of the human face. This approach was based on blind source separation of the color channels into independent components.

First, an automated face tracker was used to detect faces within the video frames which serve as the measurement region of interest (ROI) for each frame. The face detection algorithm developed by Viola and Jones, as well as Lienhart and Maydt, was used in this work. It involves a

cascade of boosted classifiers which uses Haar-like image features pre-trained from an enormous set of positive and negative images. When a face is detected, a box is defined around a face automatically and a central section of the box is selected as the ROI. The ROI is then separated into three RGB channels and spatially averaged over all pixels to yield a red, blue and green measurement point for each frame and form raw traces $x_1(t)$, $x_2(t)$ and $x_3(t)$ for each channel. The traces were subsequently processed using a 30 s moving window with 1 s increment. The raw RGB traces are normalized and then decomposed into three independent source signals using ICA. The joint approximate diagonalization of eigenmatrices (JADE) algorithm by Cardoso [10] is used to approximate statistical independence of the sources. JADE uses fourth-order cumulant tensors and involves the joint diagonalization of cumulant matrices. From the experiments, it was noticed that the second component, corresponding to the green channel, reportedly contained a strong plethymographic signal (PPG) and was selected as the desired source; otherwise the components were not ordered. The power spectrum was obtained by using the fast Fourier transform (FFT) on the selected source and the frequency at which highest power is achieved was used as the corresponding pulse frequency. The highest power was picked within an operational frequency band of 0.75-4 Hz corresponding to 45-240 bpm. The reference heart rate measurements from the recorded finger BVP signal was done in a similar way. Results from both the techniques were recorded and compared. 12 participants were tested in the experiment and a correlation coefficient of 0.98 while the *root mean squared error* (RMSE) of 2.29 was obtained for the measurement of heart rate.

2.3 Advancements in non-contact, multiparameter physiological measurements using a webcam [3]

In this work, the automated computation of heart rate from color video recordings of the human face in [2] is extended to extract multiple physiological parameters such as heart rate (HR), respiration

rate (RR) and heart rate variability (HRV). The blood volume pulse from the facial regions of a person is extracted by applying independent component analysis on the color channels.

The raw traces formed by spatially averaging over all pixels for each frame were detrended using a smoothness priors approach [11] with the smoothing parameter $\lambda = 10$ corresponding to the cut off frequency of 0.89 Hz and normalized to zero mean and unit variance. This normalized traces are then decomposed into three independent components using JADE algorithm based ICA. Once again, the order in which it returns the independent components is random. To solve this sorting problem in ICA, the component for which the peak frequency has highest power is selected in this work. Same dataset of 12 videos as used in [2] were used to obtain a correlation coefficient of 1.0 and a *root-mean-squared error* of 1.24 bpm for heart rate measurements.

The separated source signal is processed further to refine the BVP peaks using a moving average filter and a band-pass filter. This processed signal is interpolated to a frequency similar to the FDA approved Flexcom sensor used to compare the respective results. The BVP peaks were detected to calculate the inter beat intervals (IBI). Ectopic beats or motion artifacts in the signal were taken into account by using a filter on the extracted IBIs. IBI mean over the time series is used to determine the heart rate using $60/\overline{IBI}$.

2.4 Temporally constrained ICA-based fetal ECG separation [4]

In this work, cICA was successfully used to extract an underlying signal of interest. Here, a weak signal of a fetal ElectroCardioGram (ECG) was extracted by removing the dominant maternal signal.

In this method, using a prior knowledge of dominant maternal beat positions a pulse signal defined as the reference signal was generated whenever a QRS complex occurred. Peak detection algorithm was used to compute the maternal beats positions. This reference signals constraints the

cICA output to be similar and hence extracts a maternal ECG dominant signal free from both artifacts and foetal ECG. In the end, the output of cICA is eliminated from one of the abdominal ECGs that is most related to the output, to obtain the fetal ECG. This technique is depicted in the form of block diagram below.

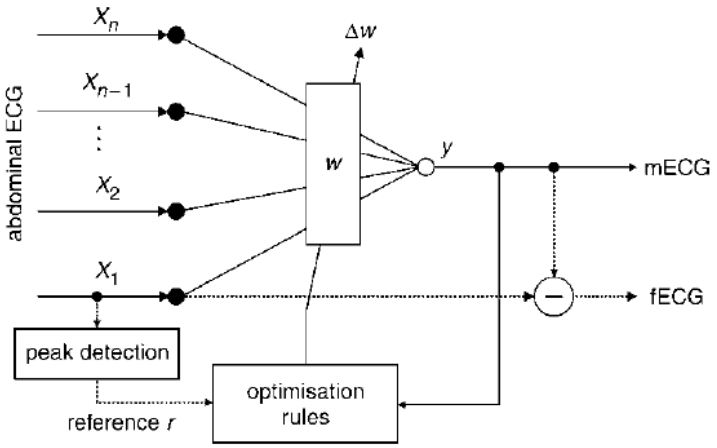


Figure 3: Block diagram of temporally constrained ICA based fetal ECG separation [4]

3 Proposed Algorithm

In this contribution, the constrained Independent Component Analysis (cICA) algorithm is used to improve the accuracy of BVP measurements using a webcam. For the pulse rate measurement, cICA can be used to help avoid the sorting problem of ICA discussed above and extract the BVP as the source of interest even with the presence of other sources.

In this work, a cICA based algorithm is designed, implemented and tested for accurately measuring the pulse rate of a person. The performance of both ICA based and cICA based pulse measurement are compared and it is showed how the latter works better. Moreover, a comparative study between both the algorithms is performed over a larger dataset of 45 videos.

A functional block scheme of the proposed algorithm is depicted in the following figure.

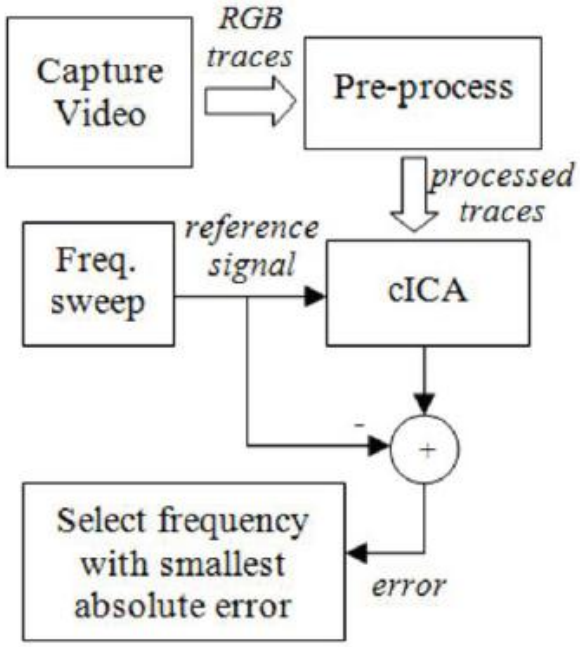


Figure 4: Functional block of the proposed algorithm

A video is captured using an RGB webcam along with face tracking for locating the entire face region as the *Region of Interest* (RoI) for pulse detection. The RGB traces from the RoI are pre-processed per trace by band-limiting to the expected heart rate range (45-240 bpm), whitened to uncorrelate the traces by linear transformation, shifted to zero mean and normalizing to unit variance to simplify the separation. For example, observed signal x is the fig. 5 (mixing of two signals s_1 and s_2 is considered here) is assumed to be already *whitened* and *normalized*. Then the problem is reduced to find an orthonormal matrix W such that components of $y = Wx$ are as independent as possible.

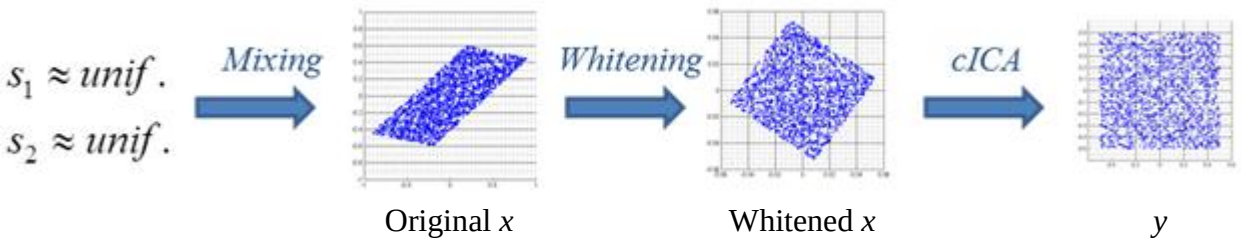


Figure 5: Demixing using cICA

The three processed traces are fed as input signals to cICA optimization as described in (29)-(30). The constrained optimization problem is solved using Newton-like learning [5] to incorporate prior information and additional requirements in the form of constraints. The reference signal fed to the cICA algorithm is a single tone with frequency corresponding to a possible heart rate in beats per minute (bpm). The peak frequency in the spectrum of the cICA output is compared to the frequency of the reference signal and the error is recorded as the absolute difference in frequencies.

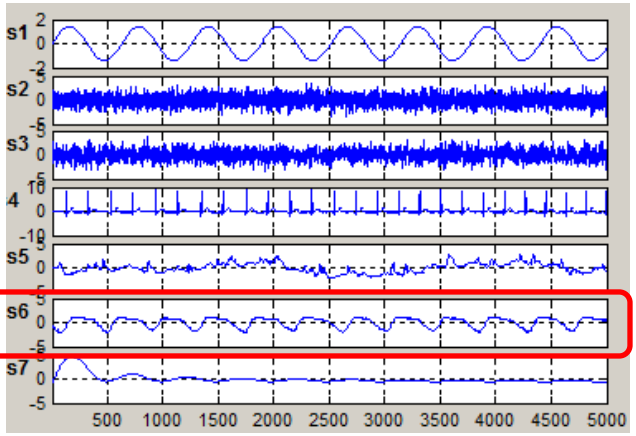
This process is repeated while sweeping through frequencies of the reference signal within the expected range starting from 0.75 Hz to 4 Hz (which provides a wide pulse range of 45 to 240 bmp) with an interval of $[f_s/(2 * N)]$ where, f_s = number of frames per second, N = total number of frames in the video sample. Note that this interval is changeable depending on the desired precision for the pulse rate. In this case it is 15/(2*900) = 0.00833 Hz or 0.5 beats per minute.

Here the operational range of pulse frequency, which is 0.75-4 Hz, is used as a priori information for the system. For each different reference signals the optimization algorithm is executed which tries to converge the one-unit constrained source separation output towards it. The reference frequency which results in the minimal absolute error is selected as the estimated pulse rate.

3.1 cICA

cICA is a useful extension to ICA which incorporates more assumptions and prior information available on underlying sources, avoiding local minima and increasing the quality of separation. In this optimization process, a single source close to the rough template or the reference signal available is successfully extracted (refer fig. 6). Thus, cICA is a more systematic and flexible approach that solves the ill-posed sorting problem of ICA.

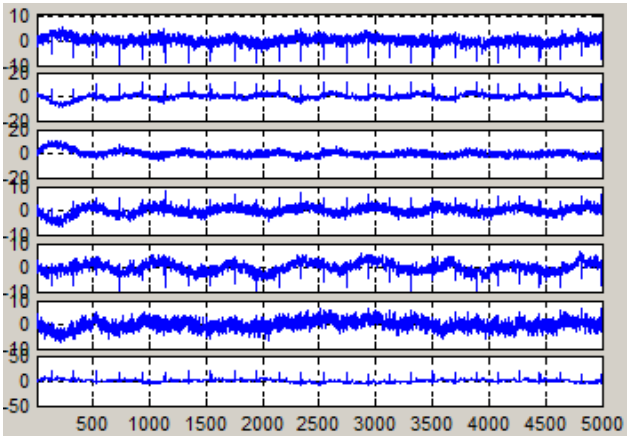
Source: s



Particularly interested in this IC

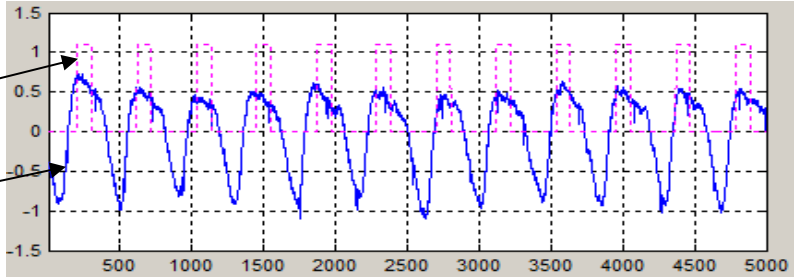
Mixing

Observation: x



Constrained ICA

Separated Signal: y



Reference Signal

Desired IC was successfully separated

Figure 6: Constrained ICA

The cICA constrained minimization problem [5] is framed as:

Minimize: $\mathcal{C}(y)$ (3)

Subject to: $g(y) \leq 0 \; and \; h(y) = 0$ (4)

Where $\mathcal{C}(y)$ denotes cICA contrast function, and $g(y)$ and $h(y)$ are inequality and equality constraints, respectively. Lagrange multiplier is used to optimize the above equation.

Augmented Lagrangian function $\mathcal{L}$ is given by:

$$\mathcal{L}(W, \mu, \lambda, z) = \mathcal{C}(y) + \lambda\, h(y) + \frac{1}{2}\,\gamma\, \|h(y)\|^2 + \mu\, \hat{g}(y) + \frac{1}{2}\,\gamma\, \|\hat{g}(y)\|^2 \quad (5)$$

$$\text{where } \hat{g}(y) = g(y) + z^2, \quad (6)$$

is used to transform the inequality constraints to equality constraints by introducing slack variable z. μ and λ are positive lagrange multipliers for inequality and equality constraints, respectively. γ is the penalty parameter which is always positive and $\|.\|$ is the Euclidean norm.

Elimination of slack variable 'z' according to ref [6]:

Substituting $\hat{g}(y)$ in equ (5),

$$\mathcal{L}(W, \mu, \lambda, z) = \mathcal{C}(y) + \lambda\, h(y) + \frac{1}{2}\,\gamma\, \|h(y)\|^2 + \mu\, [g(y) + z^2] + \frac{1}{2}\,\gamma\, \|g(y) + z^2\|^2 \quad (7)$$

$\mathcal{L}$ needs to be minimized w.r.t. (W, z) for various values of μ, λ and γ.

$$\min \mathcal{L}(W, \mu, z) = \underbrace{C(y) + \lambda\, h(y) + \frac{1}{2}\,\gamma\, \|h(y)\|^2}_{independent\ of\ z} + \min\, [\mu\, [g(y) + z^2] + \frac{1}{2}\,\gamma\, \|g(y) + z^2\|^2] \quad (8)$$

$$\text{let, } k = z^2 \text{ and s} = [\mu\, [g(y) + k] + \frac{1}{2}\,\gamma\, \|g(y) + k\|^2] \quad (9)$$

To get to its global minimum, find the derivative of s and equate it to zero.

$$\frac{ds}{dk} = \mu + \gamma \left\| g(y) + \hat{k} \right\| = 0 \quad (10)$$

$$=> \hat{k} = -\left[\frac{\mu}{\gamma} + g(y)\right] \quad (11)$$

Checking for minima/maxima: $\frac{d^2 s}{d k^2} = \gamma > 0$ (12)

So, $\hat{k}$ is the solution to minimize 's'

$$\text{Since, } k = z^2 => \hat{k} \geq 0 \quad (13)$$

$$\text{Therefore, } \hat{k} \text{ can have two values i.e., } k = \max\left\{0, -\left[\frac{\mu}{\gamma} + g(y)\right]\right\} \quad (14)$$

Now, substituting k back in $\hat{g}(y)$: $\hat{g}(y) = g(y) + z^2 = g(y) + k$

$$\hat{g}(y) = g(y) + \max\left\{0, -\left[\frac{\mu}{\gamma} + g(y)\right]\right\}$$

$$\text{Denote, } g(W, \mu, \gamma) = \hat{g}(y) = max\{g(y), -\mu/\gamma\} \quad (15)$$

$$\mathcal{L}(W, \mu, \lambda) = \mathcal{C}(y) + \underbrace{\lambda\, h(y) + 1/2\ \gamma \left\| h(y) \right\|\hat{}2}_{\mathcal{H}(y)} + \underbrace{\mu\, g(W, \mu, \gamma) + 1/2\ \gamma \left\| g(W, \mu, \gamma) \right\|\hat{}2}_{\mathcal{G}(y)} \quad (16)$$

$$\mathcal{G}(y) = max\ \{\mu g(y), -\mu^2/\gamma\} + \frac{1}{2}\gamma \| max\{g(y), -\mu/\gamma\} \|^2$$

$$= \frac{1}{2\gamma}\ [\max\ \{2\mu\gamma\ g(y), -2\mu^2\} + \ \|\max\ \{\gamma g(y), -\mu\}\|^2]$$

$$= \frac{1}{2\gamma}\ [2\mu \max\ \{\gamma\ g(y), -\mu\} + \ \|\max\{\gamma g(y), -\mu\}\|^2]$$

$$= \frac{1}{2\gamma}\ [2\mu \max\ \{\mu + \gamma\ g(y), 0\} - 2\mu^2 + \ \|\max\{\mu + \gamma g(y), 0\} - \mu\|^2]$$

Let, $l = \max\{\mu + \gamma g(y), 0\}$

$$\mathcal{G}(y) = \frac{1}{2\gamma}\,[2\mu l \;-\; 2\mu^2 + \;\|l - \;\mu\|^2]$$

$$= \frac{1}{2\gamma}\,[2\mu l \;-\; 2\mu^2 + \;l^2 + \;\mu^2 - \;2\mu l]$$

$$= \frac{1}{2\gamma}\,[l^2 - \;\mu^2]$$

Therefore, $\mathcal{G}(y) = \frac{1}{2\gamma}\,[\{\max\{\mu + \gamma g(y), 0\}\}^2 - \;\mu^2\,]$ (17)

To summarize equation (5) after slack variable elimination:

$$\mathcal{L}(W, \mu, \lambda) = \;\mathcal{C}(y) + \;H(y) + \;\mathcal{G}(y) \tag{18}$$

where, $H(y) = \;\lambda\, h(y) + \frac{1}{2}\,\gamma\,\|h(y)\|^2$ (19)

and $\mathcal{G}(y) = \frac{1}{2\gamma}\,[\{\max\{\mu + \gamma g(y), 0\}\}^2 - \;\mu^2\,]$ (20)

The gradient of $\mathcal{L}(W, \mu, \lambda)$ is obtained by partially differentiating the objective function in equ (18):

$$\nabla_W \mathcal{L}(W, \mu, \lambda) = \;\nabla_W \mathcal{C}(y) + \;\nabla_W H(y) + \;\nabla_W \mathcal{G}(y) \tag{21}$$

By partially differentiating above equation w.r.t. W, the Hessian matrix $\nabla_w^2 L(W, \mu, \lambda)$ can be obtained. Newton-like algorithm is used to optimize W:

$$\Delta\, Vec(W) = \;-\eta\,(\nabla_w^2 L)^{-1}\,\nabla_W \mathcal{L} \tag{22}$$

where η is gradually decreased for stable convergence. Gradient ascent method is used to update μ and λ as follows:

$$\Delta\mu = \;\gamma\hat{g}(y) = \;\gamma g(W, \mu, \lambda) \;=\; \gamma \max[g(y), -\,{}^{\mu}/_{\gamma}]$$

$$\Delta\mu = \;\max[\gamma g(y), -\mu] \tag{23}$$

$$\text{and, } \Delta\lambda = \gamma h(y) \tag{24}$$

Initialization: W is chosen as a uniform random vector, and μ and λ as zeros. With each iteration vector W, μ and λ are updated using equations (22), (23) and (24) until convergence.

The parameters are selected and suitably adjusted to maintain the independence criteria and the constraint properties valid every time. The choice of the contrast function and the constraints used affects the algorithms consistency. The contrast function is chosen as any function whose optimization enables the estimation of independent components.

Negentropy is used to separate the independent components. The negentropy of a signal is given by,

$$J(y) = H(y_G) - H(y) \tag{25}$$

Where y_G is a gaussian random variable with same variance as the signal y. Since the gaussian signals have the maximum entropy of zero, the negentropy is always nonnegative. The contrast function of cICA is given by:

$$\mathcal{C}(y) = -J(y) \tag{26}$$

Large $J(y)$ means that y is far from Gaussian. Since the source distribution is not known, the value of $J(y)$ cannot be computed from its definition. $J(y)$ has to be estimated from observed data: $y(t) = W\,x(t), t = 1, \dots, T$.

So, J is estimated using an approximation described in ref [7] as follows:

$$J(y) \approx \rho\,(E\{f(y)\} - E\{f(v)\})^2 \tag{27}$$

Where, ρ is a positive constant, f is a non-quadratic function and v is a zero mean unit variance gaussian variable. $f(.)$ can be any practical non-quadratic function to approximate the neg-entropies of the sub-Gaussian or super-Gaussian signals.

$$J(y) = kurt(y)^2$$

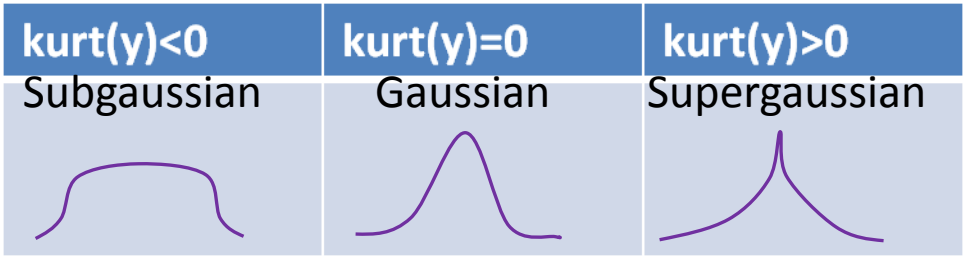


Figure 7: Representation of Sub-Gaussian, Gaussian and Super-Gaussian signals.

e.g., $f_{sub}(y) = y^4$

$f_{sup}(y) = \log\cosh y$

f is chosen depending on whether our source signal of interest is super-gaussian or sub-gaussian. In real time, most of the signals are super-gaussian nature and so we set f as proposed in ref [7] as:

$$f(y) = a^{-1}\log(\cosh(ay)) - 0.5ay^2, \tag{28}$$

where $a \in R^+$. This $f(y)$ helps achieve a stable global solution at the global optima as its second derivative is negative. Optimization of (26) alone may cause different outputs estimate the same independent component. Therefore, the uncorrelation constraint is introduced to prevent that. The equality constraint in (4) is introduced as $h(y) = (E(y^2) - 1)^2$ to restrict each output to have unit variance while the inequality constraint in (4), $g(y)$ is introduced to guide the separation of ICs. $g(y)$ incorporates additional information available on the interested signals as a priori constraints. Traces of the interesting sources, referred to as the reference signals which are not identical to the corresponding sources, carry some information to separate the desired components. These constraints the ICA algorithm to extract ICs that are closest to the corresponding reference signals. Fig (6) shows an example where a specific independent component is extracted from its mixture whose rough template is available a priori:

Choosing g(y) according to [7]: Some norm, $\varepsilon(y,r)$, can be used to measure the closeness between the estimated output y and the corresponding reference signal r. The minimum value of the norm indicates that the estimated output is the desired IC which is closest to reference signal. Now, if the desired IC is the one and only closest to the reference r, $\varepsilon(y*,r) < \varepsilon(yo,r)$ where y* is the output closest to r and y^o is the output with the next closest value of the norm. Thus, a constraint can be defined for the desired output component to have the closeness measure with r, which is less than or equal to a threshold parameter ξ. Thus, $g(y) = \varepsilon(y,r) - \xi \leq 0$ and when $y = y^*$ and none of the other (two in this case) sources corresponds to the reference r if the threshold ξ is chosen in the scalar range $\Upsilon = [\varepsilon(y^*,r), \varepsilon(y^0,r))$. The Mean Square Error (MSE) is used as a measure of closeness.

Initially, ξ is chosen to be very small so as to avoid any local minima. It is then gradually increased within its range to converge to the global minima. To summarize, the overall cICA model looks as follows:

Minimize:

$$\mathcal{C}(y) = -\mathcal{J}(y) \approx -\rho(\mathrm{E}\{f(y)\}-\mathrm{E}(f(v)\})^2 \tag{29}$$

Subject to:

$$h(y) = (\mathrm{E}\{y^2\}-1)^2 = 0 \; ; \; g(y) = \varepsilon(y,r) - \xi \leq 0, \tag{30}$$

Substituting C(y), H(y) and G(y) in equ (21):

$$\nabla_W \mathcal{L}(W,\mu,\lambda) = E\left[\nabla_y \mathcal{C}(y)x^T\right] + 2\nabla_{\sum_{yy}} \mathcal{H}(y)E\{yx^T\} + E[\nabla_y \mathcal{G}(y)x^T] \tag{31}$$

From (1): $$\nabla_y \mathcal{C}(y) = -2\rho(\mathrm{E}\{f(y)\} - \mathrm{E}(f(v)\})f'_{\mathrm{y}}(\mathrm{y})$$

$$= -\hat{\rho}f'(y) \tag{32}$$

where, $$\hat{\rho} = 2\rho(\mathrm{E}\{f(y)\} - \mathrm{E}(f(v))\}) \quad (32a)$$

From (): $$\nabla_{\sum_{yy}} \mathcal{H}(y) = \lambda h'(y) = 2\lambda(\mathrm{E}\{y^2\} - 1) \quad (33)$$

And, $$\nabla_y \mathcal{G}(y) = \mu g'(y) \quad (34)$$

To get to $\Delta\, Vec(W)$, $(\nabla_w^2 L)^{-1}$ should be known according to equ (22). In order to avoid matrix inversion, the Hessian matrix $\nabla_w^2 L$ is approximated as:

$$\nabla_w^2 L(W, \mu, \lambda) \cong D \otimes \sum_{xx} \quad (35)$$

where, $$D = d(y, \mu, \lambda) = 4\lambda(3\underbrace{E\{y\text{^}2\}}_{equals\ 1} - 1) - E\{\hat{\rho} f''(y)\} + E\{\mu g''(y)\} \quad (36)$$

or, $$D = d = 8\lambda - E\{\hat{\rho} f''(y)\} + E\{\mu g''(y)\}$$

and, $\sum_{xx}$ = Covariance matrix of input x

Using Newton-like algorithm from equ (22):

$$\Delta\, Vec(W) = -\eta\, (\nabla_w^2 L)^{-1}\, \nabla_W \mathcal{L}$$

$$= -\eta\, (D \otimes \sum_{xx})^{-1}\, \nabla_W \mathcal{L} \quad (37)$$

Applying the following relational properties:

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \quad \text{I}$$

and, $$(C^T \otimes A) Vec(B) = Vec(ABC) \quad \text{II}$$

Using I: $$\Delta Vec(W) = -\eta (D^{-1} \otimes \sum_{xx}^{-1})\, \nabla_W \mathcal{L} \quad (38)$$

Let, $\sum_{xx}^{-1} = A$, $\nabla_W \mathcal{L} = B$ and $D^{-1} = C$

Using II: $$\Delta Vec(W) = -\eta Vec(\sum_{xx}^{-1}\, \nabla_W \mathcal{L}(W, \mu, \lambda))\, D^{-1} \quad (39)$$

Transforming the vector form back to matrix form:

$$\Delta W = -\eta(\nabla_W \mathcal{L}(W,\mu,\lambda)/\ D)\ \sum_{xx}^{-1} \tag{40}$$

(Note: D is a scalar here).

or, $$\Delta W = -\eta((E\left[\nabla_y \mathcal{C}(y)x^T\right] + 2\nabla_{\sum_{yy}} \mathcal{H}(y)E\{yx^T\} + E\left[\nabla_y \mathcal{G}(y)x^T\right])/\ D)\ \sum_{xx}^{-1} \tag{41}$$

or, $$\Delta W = -\eta(\Phi(y,\mu,\lambda) + \ \Omega(y,\mu,\lambda) + \ \Psi(y,\mu,\lambda))\ \sum_{xx}^{-1} \tag{42}$$

with, $$\Phi(y,\mu,\lambda) = E\left[\nabla_y \mathcal{C}(y)x^T\right]/\ D, \tag{43}$$

$$\Omega(y,\mu,\lambda) = [2\nabla_{\sum_{yy}} \mathcal{H}(y)E\{yx^T\}]/D\ , \tag{44}$$

$$\Psi(y,\mu,\lambda)) = \ E\left[\nabla_y \mathcal{G}(y)x^T\right]/\ D \tag{45}$$

while updating μ and λ as in (23) and (24).

3.2 Stability Analysis

The suitable selection of the parameters of the cICA and the convergence stability is discussed in this section. With Newton-like learning, the algorithm is able to reach the minimum of the objective function and produce the optimum output y^* to extract the desired ICs. The threshold ξ decides the convergence of the cICA algorithm so it has to be suitably selected in the range ϒ discussed above. Any component c other than the desired source with the closeness measure $\varepsilon(c,r) \leq \xi$ corresponds to the local minima. One and only one desired IC is obtained as the optimum output at the global minimum if ξ is selected suitably in the range ϒ. However, if ξ is selected too small, the corresponding constraint $g(y) \gg 0$ causes the learning unpredictable and hence no desired IC is produced. On the other hand, if ξ is selected above the upper bound of ϒ a different IC is produced. In real-time, ξ is selected to be small in the algorithm to avoid any local minima and is gradually

increased according to the range ϒ to converge at the global minima resulting in one and only one desired IC.

The stability of the algorithm at the global optima is examined by testing if the Hessian matrix $\nabla_w^2 L(W, \mu, \lambda)$ is positive-definite. When the Hessian matrix is positive definite, the cICA algorithm converges stably. The approximated Hessian matrix in (35) is always positive-definite when the input covariance matrix $\sum_{xx}$ is non-singular and the element D is positive. The former condition is true in most of the cases when a large number of samples of the signal are available. Considering the first term $d_{\mathcal{H}}^*(y^*, \lambda^*) = 4\lambda^*(3E\{y^{*2}\} - 1)$ from (36): since the variance of y* approaches to one, $d_{\mathcal{H}}^*(y^*, \lambda^*) = 8\lambda^*$ and as λ^* is always positive the first term in (36) is also positive. Now, considering the second term $d_j^*(y^*) = -E\{\hat{\rho}^* f''(y)\}$: The positivity of $d_j^*(y^*)$ is subject to Gaussianity of the signals. From the non-linear function $f(y)$ proposed in (28), it can be shown that $-f''(y) \geq 0$ and the value $\hat{\rho} = 2\rho(\mathrm{E}\{f(y)\} - \mathrm{E}(f(v))\}$ in (32a) is always positive for the optimal solution with the super-gaussian distribution. Therefore, the product $-\hat{\rho} f'(y)$ is positive for all super-gaussian signals. Considering the third term $\mu^* g''(y*)$ which corresponds to the closeness between output and reference signals: In the inequality constraint, it is important to select a suitable distance as a closeness measure $\varepsilon(y, r)$. The Mean Square Error (MSE),$\varepsilon(y, r) = E\{(y - r)^2\}$, being most simple and common measure can be used as a distance measure between estimated desired IC 'y' and the corresponding reference signal 'r'. It is required that both 'y' and 'r' should be normalized to same mean and variance. Correlation can also be used, $\varepsilon(y, r) = 1/(E\{yr\})^2$, with both the output signal and the reference signal normalized so that the correlation is bounded. Depending on the form of the reference signal available, the selection of the closeness function can be different from one output to another. It can be seen that using any of the closeness measure discussed above, the term $\mu^* g''(y*)$ is always positive for positive Lagrange multipliers μs. Therefore, when the non-linear function $f(y)$ in (28) is used for super-gaussian signals and MSE or

correlation as the closeness measure, the convergence of the learning algorithms in (42) is always stable.

4 Experimental Study

One minute color video recordings were made using a 2 Mega Pixel RGB Logitech camera at 15 frames per second with pixel resolution of 320x240 and saved in WMV/AVI format. A total of 45 participants between the ages of 18-45 years of both genders with varying skin colors and from various nationalities were enrolled in this study. The study was approved by the Internal Review Committee for Protecting Human Subjects at the Rochester Institute of Technology. An informed consent form was signed by participants prior to the start of each study session. During the video recording, participants were asked to sit still in a relaxed state. At the same time, a reference heart rate or a ground truth was measured using a BVP finger probe oximeter for validation. All videos were further processed offline using MATLAB.

First, an automatic face tracker was used to detect the faces and locate the ROI within each video frame. Free Face Detection Toolbox version 0.21 using Local Binary Patterns and Haar features [9] were utilized. The Face Detection algorithm was mainly written in C and wrapped with a MATLAB interface. The ROI was then separated into three RGB channels by averaging all the pixels in a frame to obtain a single red, blue and green measurement point for each video frame and form the raw traces. Each video resulted in a block of three vectors of size 900 (30 sec x 15 fps) and is combined to form a 2-dimension matrix of size 3x900, with each row representing a channel. These raw traces were pre-processed by first using a band-pass filter of 0.75-4 Hz, assuming that the pulse lies within [45-240] bpm for a normal human, then whitened to uncorrelated the three sources and normalized to unit variance for simplifying the separation. The normalized traces were fed to the cICA algorithm with reference signal sweep, within the expected heart rate range of 0.75-4 Hz, as described above to extract the BVP with a sweep resolution of 0.5 bpm. The MATLAB

periodogram function was used to extract the peak frequency with a hamming window to smoothen the signal for power spectral density estimation. All video streams were also used for cross validation with the ICA pulse rate measurement algorithms described in [2, 3]. The BVP measurements using both ICA techniques in [2, 3] were compared against state-of-the-art cICA algorithm to estimate the corresponding performances. An example of a video frame showing ROI selected using automated face tracker with raw traces extracted is shown in the figure below:

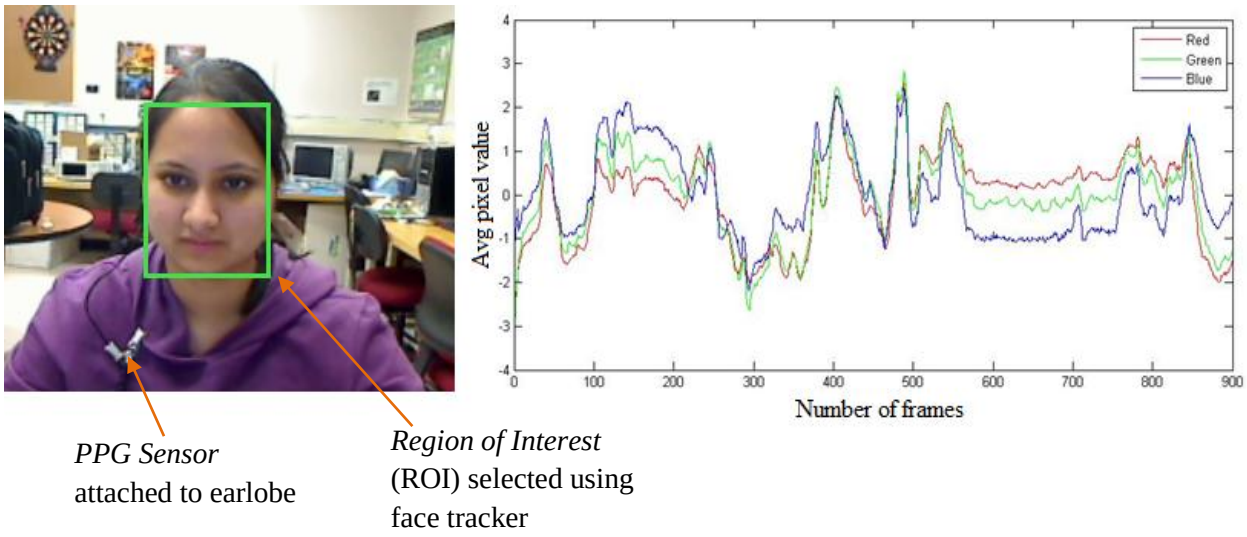


Figure 8: (a) Video frame showing ROI and PPG sensor, (b) Signals from RGB channels.

5 Measuring Similarity

For combined graphical and statistical measurements of different pulse measurement techniques - Bland-Altman plots can be used. The differences between the estimated pulse rate from ICA/cICA algorithm and the BVP finger sensor are plotted against the averages of both the systems. The root mean squared error (RMSE) was also calculated to determine the performance of each system.

6 Results and Discussion

A representative video is selected to analyze the results. The error (in Hz) as a function of the reference signal frequency in bpm is depicted in figure 9. It can be observed from this plot that the

error for a single frequency, very close to the true pulse rate of 84 bpm (in this case), is almost zero and the error increases linearly as one move away from the True Pulse. Same behavior was observed for all other videos in the dataset. This implies that the proposed algorithm does not suffer from ambiguities associated with convergence to local minima.

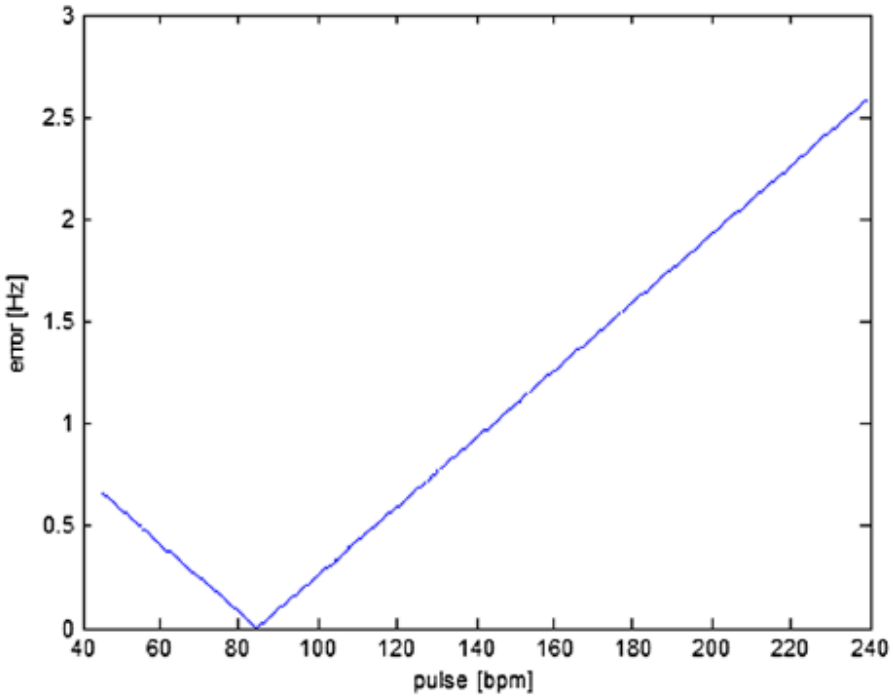


Figure 9: Error vs. frequency sweep in beats per minute (bpm).

A photoplethysmography (PPG) signal is recorded using the MP36 biopac system with a sensor attached to the earlobe of the person at the same time the one-minute video recording was made. The cICA output signal when the minimum or zero error is achieved along with its corresponding reference signal is plotted against the PPG signal obtained via the biopac system in figure 10. The close resemblance between the signals is evident.

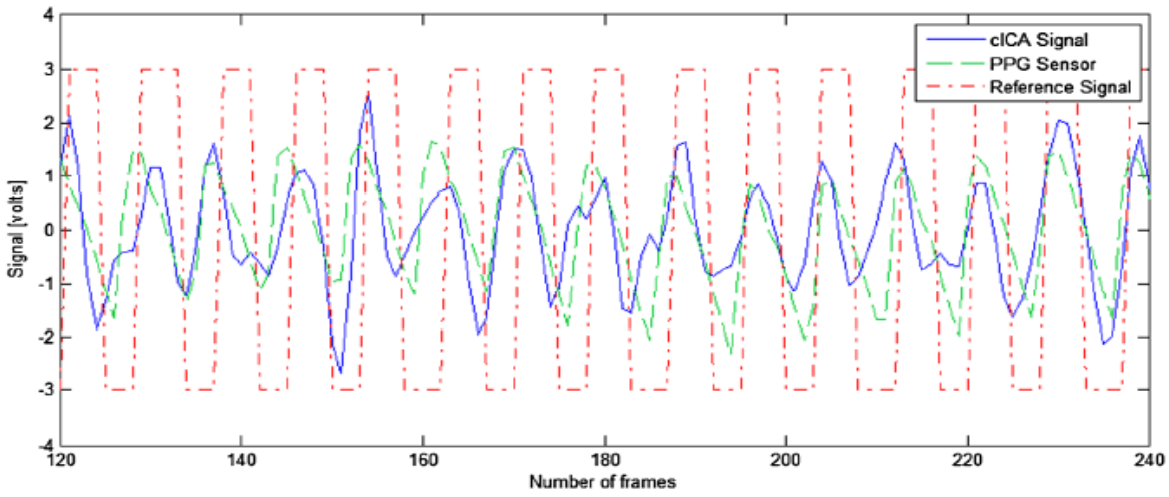


Figure 10: Photoplethysmography (PPG) sensor signal compared to cICA output signal.

Experiments were conducted at the laboratory of Rochester Institute of Technology and the sorting problem of the ICA based approach was clearly visible. The problem was resolved using the proposed cICA approach. Tab. 1 presents pulse detection results along with the heart rate measured by the finger probe oximeter (True Pulse) for a representative set of 45 videos. The detected pulse from each ICA component is presented for the ICA approach. For each video a different component provided a result closer to the True Pulse. This effect could be observed clearly from the first seven samples in the table. Moreover, the best ICA component is selected manually and appended in this table. The cICA algorithm provided a similar result in accuracy to the best result across all ICA components for most of the videos. Thus, with cICA algorithm there is no ambiguity regarding component selection and the estimated output is confirmed to be the desired pulse signal. Note that for videos 8, 9, 10 and 35 cICA performed better than all ICA components. One should also note that for videos 28, 29, 32, 33 and 38 none of the algorithms worked really well. The reason behind this could be that there may be multiple source signals mixed in the observed RGB signals whose frequency lie within the same pulse rate frequency range or the movement of the subject during the recording might influence the result. Moreover, the ICA or cICA model is assumed to be linear while the physiological changes in the blood volume due to motion could be non-linear.

TABLE I
PULSE RATE SAMPLES FROM DATA SET IN [BPM]

Video no	True Pulse	ICA component			ICA best comp	cICA
		1st	2nd	3rd		
1	60	49.0	58.0	45.9	58	58.6
2	64	48.5	82.4	64.6	65	64.1
3	72	71.6	70.5	51.0	72	71.6
4	66	73.2	72.4	65.4	65	65.6
5	76	81.1	45.5	81.7	82	81.1
6	113	45.0	47.7	113.2	113	113.6
7	49	49.4	56.4	49.6	49	50.1
8	68	51.3	112.1	50.6	51	69.6
9	78	45.7	66.3	60.1	66	72.3
10	66	53.4	49.4	51.0	53	69.6
11	84	65.0	60.4	84.3	84	84.7
12	53	54.2	54.9	54.7	55	54.4
13	63	62.6	69.6	63	63	63.1
14	60	62.4	62.4	62.1	62	60.6
15	83	62.6	81.2	62.6	81	81.7
16	108	45.9	45.0	108.0	108	108.2
17	61	48.5	58.8	60.6	61	61.0
18	78	46.0	101.8	80.5	81	78.1
19	60	60.0	60.4	56.7	60	60.6
20	94	73.7	93.8	93.6	94	94.1
21	53	45.0	53.2	53.2	53	53.6
22	69	68.5	68.7	69.2	69	69.4
23	63	66.3	60.9	62.8	63	63.1
24	60	52.1	60.1	109.7	60	59.6
25	83	45.2	82.6	82.3	83	82.9
26	74	50.7	51.8	70.0	70	76.6
27	57	53.8	56.8	73.0	57	56.6
28	73	65.6	47.3	64.3	66	62.6
29	68	60.2	60.2	57.6	60	60.1
30	64	63.7	57.0	63.0	64	63.6
31	69	45.0	45.0	68.1	68	68.6
32	68	48.3	55.6	51.6	56	56.1
33	66	45.3	45.0	72.5	73	72.6
34	100	56.5	97.1	97.1	97	97.6
35	96	45.2	55.1	56.7	57	91.1
36	63	61.9	61.9	61.7	62	62.1
37	82	64.3	53.4	83.0	83	82.1
38	67	63.5	63.2	61.0	63	74.6
39	76	78.8	79.0	76.4	76	76.6
40	97	56.7	92.9	64.6	93	96.1
41	62	49.2	63.7	61.3	61	66.6
42	70	81.7	67.6	67.9	68	68.1
43	86	53.1	84.3	84.8	85	84.6
44	57	57.3	57.5	57.5	57	57.6
45	72	48.1	70.0	69.8	70	70.1

Results for the entire data set are summarized in Fig. 11 using Bland-Altman plots discussed above for the ICA based algorithms depicted in [2,3], the impractical scenario where the ICA component providing pulse closest to True Pulse is chosen and the proposed cICA algorithm. From the comparison of the ICA algorithms of [2-3] (Fig. 11a) with the impractical a-posterior best component selection (Fig. 11b), it is clear that truly solving the sorting problem provides much better results. By comparing the result by best component selection (Fig. 11b) with the proposed cICA algorithm (Fig. 11c), it is evident that the cICA algorithm performs better than beyond solving the sorting problem. The *Root Mean Square Error* (RMSE) for the ICA algorithm across the set when selecting the second IC (ref. [2]) is 20.6 bpm and when selecting the component with maximum spectral peak (ref. [3]) is 9.5 bpm while the RMSE for the proposed cICA algorithm is 3.5 bpm. Such inaccuracy is within the expected margin of error of the True Pulse reference measurement, implying that the proposed algorithm provides measurement accuracy similar to the BVP finger probe oximeter.

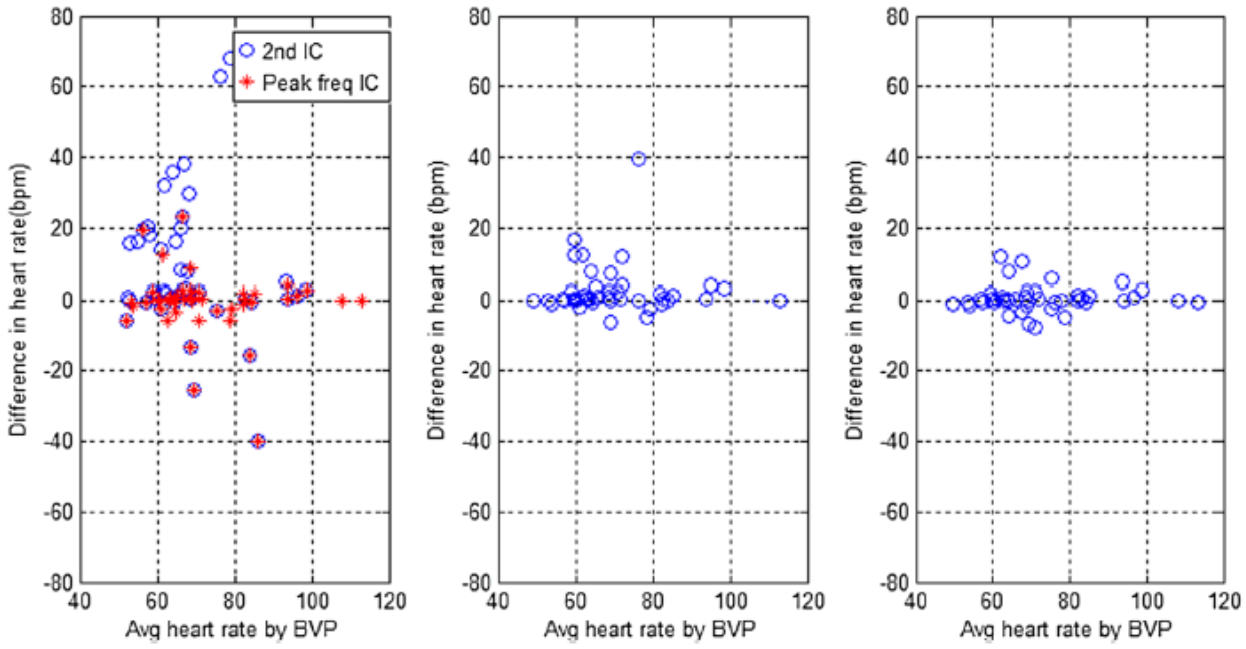


Figure 11: Bland-Altman plots: (a) independent component analysis (ICA) algorithm, (b) best component of ICA, (c) constrained ICA (cICA) algorithm.

In order to prove the robustness and consistency of the proposed algorithm, another experiment was performed in which one 10 minutes long video was captured while the subject was asked to sit in a still position. At the same time the pulse rate was recorded using the finger probe oxymeter. The time series signal obtained after running the face tracker on the video was then segmented into 30 seconds each and processed using the proposed algorithm. Tab. 2 presents and compares the results of ICA best component with the cICA based pulse estimation for this experiment. Note that for the last 6 batches, which correspond to the last three minutes of the video, the result is noisy for both the cases. This is because of the motion artifacts as noticed in the video. Otherwise, it is clearly seen that cICA performs better than ICA when we compare the results for the overall video length. Also note that for the first batch the result is off by approximately 18 beats in both the cases. The reason for this is not clear but we assume that it is because of some other dominant signal with a slightly different frequency falls within the same band as that of the pulse frequency. Further work is required to understand various artifacts and other signals with similar frequencies present in the mixture; how to separate them efficiently.

TABLE II
PULSE RATE SAMPLES FOR 10 MINUTES LONG VIDEO IN [BPM]

Batch no	True Pulse	ICA best comp	cICA
1	98	80.2	80
2	101	46.4	101
3	99	51.9	99
4	100	99.5	100
5	94	50.5	95
6	97	98.4	98
7	96	96.9	96
8	98	58.9	97
9	100	99.3	100
10	98	60.9	98
11	96	96.0	95
12	98	75.1	98
13	95	89.9	94
14	97	97.1	97

15	94	79.1	74
16	96	92.9	85
17	92	63.5	64
18	90	94.5	90
19	90	92.3	90
20	92	45.3	67

Motion artifact is one of the major limitations of this study. Irrespective of the type of algorithm incorporated, the motion of subjects during video recording affected the accuracy for remote pulse rate measurement. An attempt to overcome this limitation was made by using face tracking algorithms which was efficient in locating the face within a ROI selected for processing. However, jitter in the processed RGB traces could be observed because of the positioning of the face within the ROI would alter slightly. In order to compensate for this effect the rate at which face tracking is performed was reduced from every frame to every 15^{th} frame. In these experiments, subjects were asked to sit still therefore the effect of such motion artifacts on presented results is minimal. The primary use of this methodology is limited to be in a home environment where this pulse rate measurement technology could be incorporated in personal computers. However in a real-world application, non-obtrusive pulse rate measurement algorithms should be augmented with motion artifacts cancellation algorithms for it to work in a non-cooperative environment.

Another important limitation of cICA approach over existing ICA approach is the processing time. The cICA approach requires more processing time than the ICA approach due to the sweep performed over the reference signal. The MATLAB runtime between the two algorithms were compared and it was found that the cICA algorithm takes 30 times longer to run. Remember that this result was obtained when the sweep was defined with intervals of 0.5 bpm. However, if the sweep is performed using an interval of 1 bpm, the processing time ratio is reduced to 15. Since pulse measurement is not expected to be performed continuously, processing time is not considered

to be a limited factor in practical implementation. Moreover, if embedded programming possibilities are considered the iterations in sweeping the frequency over the band range could be performed fasted thereby reducing the ratio of the processing time between the two approaches.

7 Future Work

The ability to extract a photoplethysmography (PPG) – like signal non-obtrusively using constrained ICA approach to RGB video data has the potential to revolutionize detection of various diseases as well. The proposed technique focused on cardiac pulse detection with PPG signal extraction. In future, the energy of this PPG signal could be used to detect cancerous liver in endoscopic video streams. Moreover, non-obtrusive detection of peripheral vascular disease caused by arteriosclerosis and acute ischemia of the heart during open heart surgery could be possible by extending this work.

8 Conclusion

A novel cICA algorithm for non-obtrusive cardiac pulse rate measurement from video recordings of the human face using a webcam was proposed and implemented. The proposed methodology was evaluated using a comparative study with state-of-the-art algorithms over a large data set. The results demonstrated how the ICA sorting problem was resolved by forcing the ICA optimization process to extract a single source closely related to the reference signal. The sweep of frequencies in the reference signal was used as an additional step in the cICA algorithm. While estimating the pulse rate of a human, the results also demonstrates how a photoplethysmography (PPG) equivalent signal could be successfully extracted using this low-cost automated technique. The estimation error of the proposed algorithm was shown to be within the estimation error of a commercially available finger-probe oximeter, suggesting that non-obtrusive pulse rate measurements could replace existing practices. This technology when combined with the existing motion artifacts cancellation

algorithms is very promising for extending and improving access to medical care. Although this contribution focused on addressing pulse rate measurement, the proposed cICA algorithm is likely to be useful for estimating other physiological parameters such as respiratory rate, heart rate variability, blood oxygen saturation, etc. or where an ICA sorting problem is evident. The basic requirement of this approach is the existence of prior knowledge on the parameter being estimated.

References

[1] C. Takano, and Y. Ohta, “Heart rate measurement based on a time-lapsed image,” *Med. Eng. Phys.* 29(8), 853-857, 2007.

[2] 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*, 18(10), 10762-10774, 2010.

[3] M.Z. Poh, D.J. McDuff and R.W. Picard, "Advancements in non-contact, multiparameter physiological measurements using a webcam," *IEEE Trans. on Biomedical Engineering*, vol. 58, no. 1, Jan 2011.

[4] J. Lee, K.L. Park and K.J. Lee, “Temporally constrained ICA-based fetal ECG separation,” *Electronics Letters*, vol. 41, no. 21, Oct. 2005.

[5] W. Lu and J.C. Rajapakse, “Approach and Applications of Constrained ICA,” *IEEE Trans. on Neural Networks*, vol. 16, no. 1, Jan. 2005.

[6] D.P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, New York: Academic, 1982.

[7] J.F. Cardoso, “High-order contrasts for independent component analysis,” *Neutral Comput.* 11(1), 157-192 (1999).

[8] A. Hyvärinen, J. Karhunen and E. Oja. Independent Component Analysis. Wiley-Interscience, 1st Ed., 2001.

[9] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, “Remote plethysmographic imaging using ambient light,” *Optics Express*, 16(26), 21434–21445, 2008.

[10] Sébastien Paris , *Mathworks Face Detection Toolbox 0.21c*, Available at: http://webscripts.softpedia.com/script/Multimedia/Video/Face-Detection-Toolbox-54816.html

[11] M.P. Tarvainen, P.O. Ranta-Aho, and P.A. Karjalainen, “An advanced detrending method with application to HRV analysis,” *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp. 172-175, Feb. 2002.

ACKNOWLEDGEMENT

This work was a part of the RIT Corporate R&D project: **Image Processing for Non-Invasive Triage** (2010-2011; PI: Gill R. Tsouri, co-PI: Sohail Dianat), sponsored by the *Xerox Technology Incubation Network* (XTIN) program.