A Simple, Remote, Video Based Breathing Monitor

Nir Regev

essence
Better life made possible™

Essence
Partner of Choice Worldwide

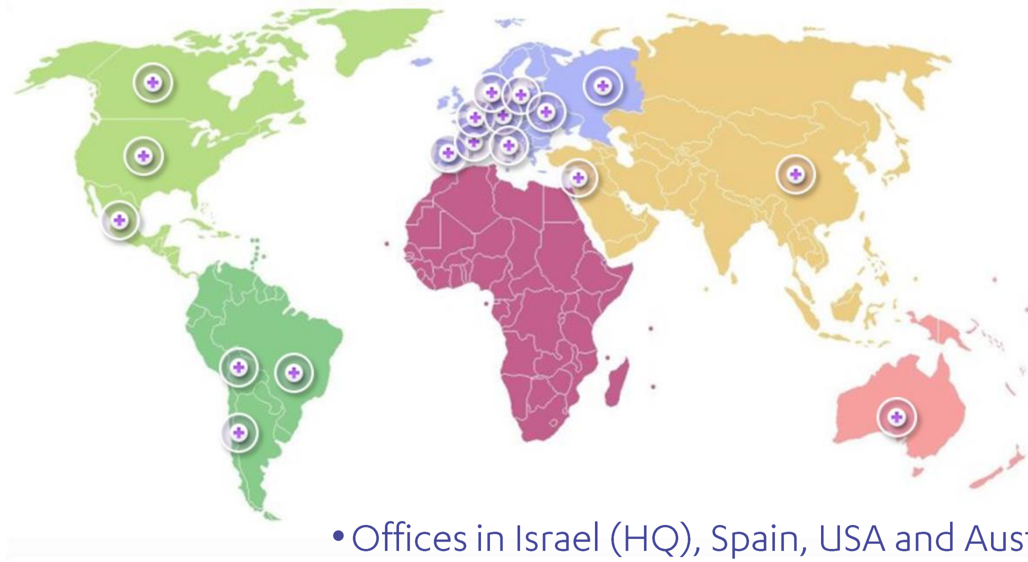
M2M **connected-living** solutions

23+ years serving global service providers

250 professionals

End-to-End: hardware, software and services

Award winning solutions



- Offices in Israel (HQ), Spain, USA and Australia
- Manufacturing lines in Israel and Hungary
- Local partners worldwide

ESX2015 INNOVATION AWARDS —WINNER—

Broadband INFOVISION AWARDS Best Home Service Innovation

European consumers choice 9.4

BUSINESS IMPACT AWARD 2015

14/15 IoT/M2M INNOVATION WORLD CUP

essence Better life made possible™

Main idea

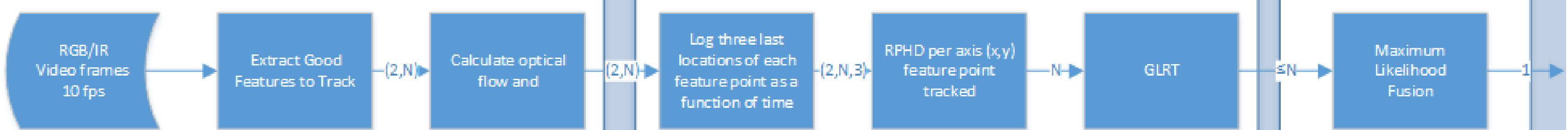
- Track the torso movement through time from the video
- Use this information to estimate the breathing frequency

essence
Better life made possible™

Motivation

- Complimentary algorithm that can “take over” when facial vid is not available
- Color based and motion based algorithms needs a “good look” at the face of the subject
- Complement other modalities
- Possible applications:
  - Elderly care
  - Baby monitoring

Algorithmic flow



Computer Vision

Video frames at 10 frames per second are acquired from the sensor. N feature points are extracted and an optical flow is calculated based on the Lucas-Kanade algorithm.

Time domain real-time spectral estimation

Last three samples of each feature point (x,y) locations are kept and the Reformed Pisarenko Harmonic Decomposition is used for fundamental frequency estimation in real time. Using the current estimate for each point a generalized likelihood ratio test is executed to determine if a sine-wave exists in the breathing frequency band (0.2-0.8Hz). This block outputs at the most N estimated breathing frequencies.

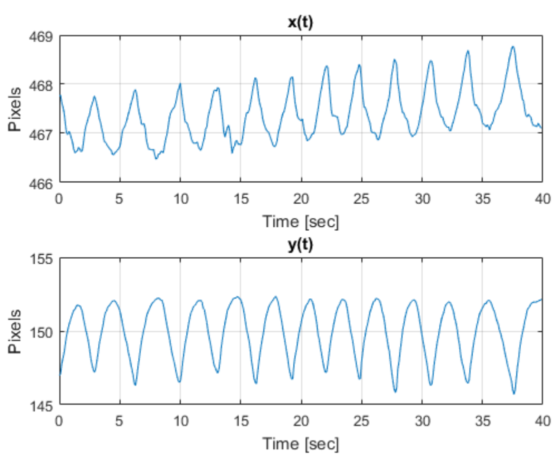
Fusion

Maximum likelihood fusion of all frequency estimates is executed to produce one optimal frequency estimate.

essence
Better life made possible™

Optical flow tracking

- Lucas-Kanade's Optical flow algorithm will track variations of points on the moving torso through time
- Variation is given both in x and the y axes
- The outcome is different in SNR, and trendy



Assumed model

- Notations:
  - N is the number of points, thus, 2N "modalities" are present (x, y)
  - $i = 0, \cdots, 2N - 1$ is the modality index
  - $k = 0, \cdots K_i - 1$ is the sample index
  - K_i is i-th modality number of observed samples

Assumed Model – cont.

$$z_i^-(k) = a_i \sin\left(\omega_b k + \varphi_{b_i}\right) + v_i^-(k),$$

- ω_b is the unknown breathing angular frequency
- φ_{b_i} is an unknown nuisance parameter
- $v_i^-(k)$ is an additive white Gaussian noise with variance σ^2 and is i.i.d. across modalities and samples.
- Note that the signal level a_i changes across modalities.

Actual Model

$$Z_i(k) = c_i k + d_i + a_i \sin\left(\omega_b k + \varphi_{b_i}\right) + v_i(k),$$

- Where c_i and d_i are the trend coefficients for modality i
- Thus, the trend needs to be iteratively estimated and removed to obtain the desired model

RLS al

- Recurs the trend-line pe
- This tr ignal
- The tr r spectral estima

Algorithm 1 Recursive Least squares line fitting

$k \leftarrow 4$ ▷ wait for the first three samples
$\alpha \leftarrow 0.99$
Construct the Vandermonde matrix $H^3 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$
Construct the observation vector $\tilde{\mathbf{z}}^3 = \begin{pmatrix} \tilde{\mathbf{z}}_0 \\ \tilde{\mathbf{z}}_1 \\ \tilde{\mathbf{z}}_2 \end{pmatrix}$
$P^k \leftarrow \left(H^{3^T} H^3\right)^{-1}$ ▷ Size: 2×2
Solve for $\boldsymbol{\theta}$: $\hat{\boldsymbol{\theta}}^k_{RLS} = P^k H^{3^T} \mathbf{z}$.
repeat
  $H^k \leftarrow (k, 1)$ ▷ Size: 1×2
  $G^k \leftarrow P^k H^{k^T} \left(H^K P^k H^{k^T} + \alpha\right)^{-1}$ ▷ Size: 2×1
  $\varepsilon \leftarrow \left(\tilde{\mathbf{z}}_k - H^k \hat{\boldsymbol{\theta}}^k_{RLS}\right)$
  $\hat{\boldsymbol{\theta}}^{k+1}_{RLS} \leftarrow \hat{\boldsymbol{\theta}}^k_{RLS} + G^k \varepsilon$
  $P^{k+1} \leftarrow \frac{1}{\alpha} \left(I_{2\times 2} - G^k H^k\right) P^k$
  $k \leftarrow k + 1$
until no more samples are available

essence
Better life made possible™

Spectral estimation

- The breathing frequency needs to be estimated real time.
- The frequency resolution should be 1BPM
- ⇒ The signal should be observed for at least 60 seconds
- We need to be faster – what to do?

Reformed Pisarenko Harmonic Decomposition

Algorithm 2 Reformed Pisarenko Harmonic Decomposition

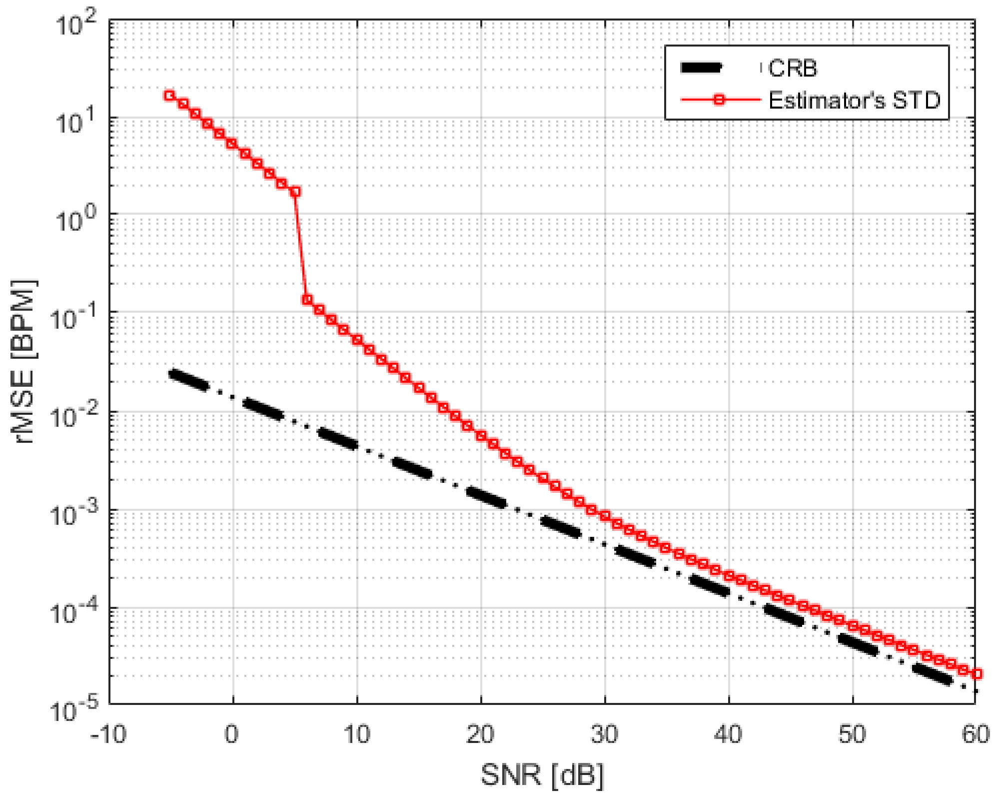
1: $k \leftarrow 2$ ▷ wait for the first three samples

$$VAR_i[\hat{\omega}_b] = \frac{\cos^2(2\omega_b) + \cos^2\omega_b}{SNR^2(K_i - 2)[2+\cos(2\omega_b)]^2\sin^2\omega_b} + \frac{1}{SNR(K_i-2)^2\sin^2\omega_b} + \frac{3 + 4\cos(2\omega_b) - cos(4\omega_b)}{4SNR^2(K_i-2)^2[2+\cos(2\omega_b)]^2\sin^2\omega_b}.$$

8: $k \leftarrow k + 1$

9: **until** no more samples are available

RPHD performance – per modality



essence
Better life made possible™

Verify the model

- All modalities with frequencies not in the breathing band are discarded
- Once a breathing frequency estimation is done, we need to verify that it adheres to our sinusoidal model
- Thus, a Generalized Likelihood Ratio Test (GLRT) is executed.
- Define two hypotheses:

$$H_0:\ z_i^-(k) = v_i^-(k)$$
$$H_1: z_i^-(k) = a_i \sin\left(\omega_b k + \varphi_{b_i}\right) + v_i^-(k),$$

Verify the model – cont.

- The GLRT is given by:

$$\frac{1}{\sqrt{K_i}}\sum_{k=0}^{K_i-1} z_i^{-}(k)[\sin\omega_b k+\cos\omega_b k] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_i$$

- γ_i is a threshold calculated using the Neyman-Pearson's theorem to keep a fixed false alarm, P_{FA}

$$\gamma_i = -\sigma^2 ln P_{FA}$$

Verify the model – cont.

- All modalities that didn't survive the GLRT are discarded
- All modalities that survived are fed into the fusing algorithm

Fusion algorithm

- At each time instance we have a refined estimation for each modality & its corresponding variance.
- Assumption: the estimation error is a zero mean Gaussian random variable with variance σ_i^2.
- Thus we have the model

$$\hat{\omega}_{b_0} = \omega_b + w_0$$
$$\hat{\omega}_{b_1} = \omega_b + w_1$$
$$\vdots$$
$$\hat{\omega}_{b_{2N_s-1}} = \omega_b + w_{2N_s-1}$$

Fusion algorithm – cont.

- $w_i \sim N\left(0, \sigma_i^2\right)$ and ω_b is the true parameter value
- Putting in a matrix form

$$\widehat{\boldsymbol{\omega}}_b = \mathbf{1}\omega_b + \boldsymbol{w}$$

- where $\boldsymbol{w} \sim N(0, R)$ and $R = diag\left(\sigma_0^2, \sigma_1^2, \cdots, \sigma_{2N_s-1}^2\right)$
- The solution of this estimation problem in the maximum likelihood sense is given by

$$\widehat{\omega}_{b_{ML}} = (\mathbf{1}^T \boldsymbol{R}^{-1} \mathbf{1})^{-1} \boldsymbol{R}^{-1} \mathbf{1}^T \widehat{\boldsymbol{\omega}}_b =$$

$$= \frac{1}{\sum_{i=0}^{2N_s-1} \frac{1}{\sigma_i^2}} \sum_{i=0}^{2N_s-1} \frac{\widehat{\omega}_{b_i}}{\sigma_i^2}$$

Experiments

- Tested on two babies and 3 adults using an RBG sensor and an IR sensor during day and night.
- The true breathing rate was counted manually from the video and was divided by the video recording time.
- The maximum error varies from 0.7 to 1 BPM after ten seconds of video.
- However, the algorithm failed in the good points to track step in a setting where the subject was wearing a pattern-less shirt.

Experiments – cont.

TABLE I

EXPERIMENTS RESULTS.

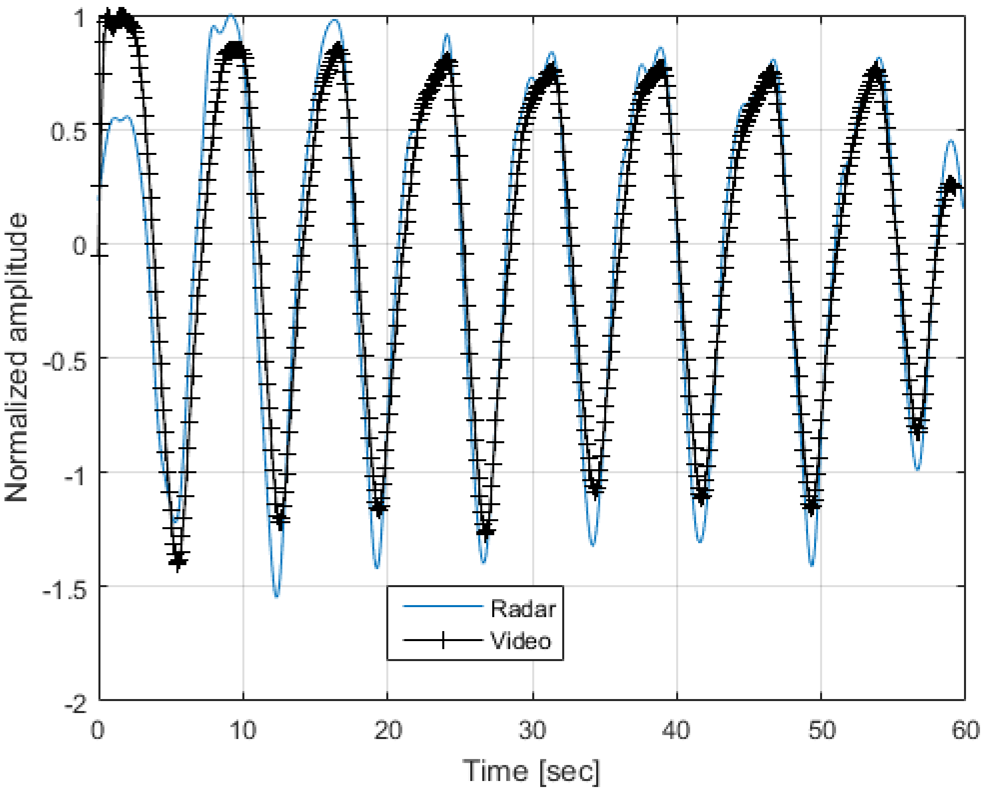
Subject gender [m/f]	Subject age [years]	Maximum error - proposed algorithm [BPM]	Maximum error - optimal algorithm [BPM]
m	40	1	0.5
m	50	0.7	0.3
m	30	1	0.4
m	2	1	1
f	0.5	1	1

DEMO - me

DEMO – my colleague

Future work

- **Sensor fusion with a radar**



essence
Better life made possible™

Thank You

www.essence-grp.com