Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions

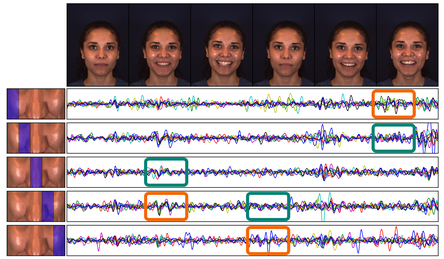
Sergey Tulyakov, Xavier Alameda-Pineda, Elisa Ricci, Lijun Yin, Jeffrey F. Cohn, Nicu Sebe

University of Trento · Fondazione Bruno Kessler · Binghamton University · Carnegie Mellon University

Contribution

- Self-adaptive matrix completion framework to discard noisy regions of the face
- Heart rate estimation under realistic conditions: lighting, facial expressions and varying head pose
- Efficient short-term heart rate measurement

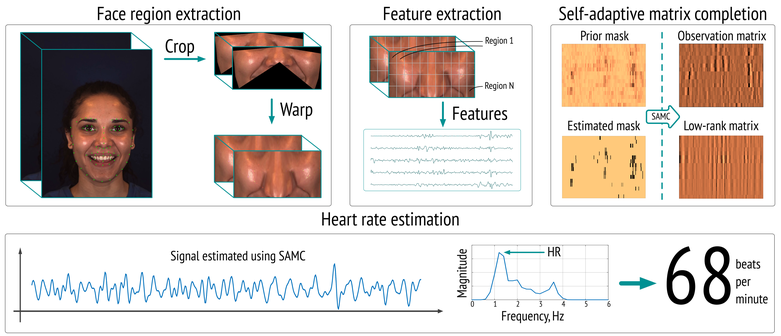
Motivation



Different parts of the face show different temporal dynamics and are subject to noise, as they are heavily affected by movements and illumination changes.

Idea: jointly estimate the HR signal and select the most reliable regions

Pipeline



Self-adaptive matrix completion

- Jointly estimates the mask and the low-rank matrix
- Automatically selects the most reliable regions
- Enforces heart rate frequency range

Formulation

$$\min_{\mathbf{E},\mathbf{M}} \nu\|\mathbf{E}\|_* + \|\underbrace{\mathbf{M}}_{\text{Mask}} \circ (\mathbf{E} - \mathbf{C})\|_{\mathcal{F}}^2 + \underbrace{\gamma\,\mathrm{Tr}(\mathbf{E}\mathbf{L}\mathbf{E}^\top)}_{\text{Temporal smoothing}} - \underbrace{\beta\|\mathbf{M}\|_1}_{\text{Anti-sparsity}} + \mu\|\mathbf{M} - \underbrace{\widetilde{\mathbf{M}}}_{\text{Mask prior}}\|_{\mathcal{F}}^2$$

Augmented Lagrangian form

$$\min_{\mathbf{E},\mathbf{F},\mathbf{M},\mathbf{Z}} \nu\|\mathbf{E}\|_* + \|\mathbf{M}\circ(\mathbf{F}-\mathbf{C})\|_{\mathcal{F}}^2 + \gamma\,\mathrm{Tr}(\mathbf{F}\mathbf{L}\mathbf{F}^\top) - \beta\|\mathbf{M}\|_1 + \mu\|\mathbf{M} - \widetilde{\mathbf{M}}\|_{\mathcal{F}}^2 + \langle \mathbf{Z}, \mathbf{E}-\mathbf{F}\rangle + \frac{\rho}{2}\|\mathbf{E}-\mathbf{F}\|_{\mathcal{F}}^2,$$

$\mathbf{E}/\mathbf{M}$ step

$$\min_{\mathbf{E}} \nu\|\mathbf{E}\|_* + \frac{\rho}{2}\|\mathbf{E} - \mathbf{F} + \rho^{-1}\mathbf{Z}\|_{\mathcal{F}}^2. \qquad \min_{\mathbf{M}} \|\mathbf{M}\circ(\mathbf{F}-\mathbf{C})\|_{\mathcal{F}}^2 - \beta\|\mathbf{M}\|_1 + \mu\|\mathbf{M} - \widetilde{\mathbf{M}}\|_{\mathcal{F}}^2,$$

Solution for $\mathbf{E}$

$$\mathbf{F} - \rho^{-1}\mathbf{Z} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$$
$$\mathbf{E}^* = \mathbf{U}\mathcal{S}_{\frac{\nu}{\rho}}(\mathbf{D})\mathbf{V}^\top$$
$$\mathcal{S}_\lambda(x) = \max(0, x - \lambda)$$

Solution for $\mathbf{M}$

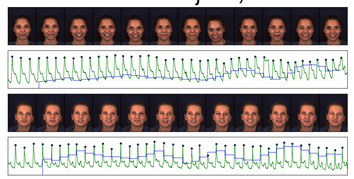
$$\min_{m_{rt}\in\{0,1\}} (f_{rt} - o_{rt})^2 m_{rt} + \mu(m_{rt} - \widetilde{m}_{rt})^2 - \beta m_{rt}$$

Datasets

Mahnob-HCI 27 subjects, 527 videos

Constrained laboratory setting
Few facial expressions and movements
Low temporal heart rate dynamics

MMSE-HR 40 subjects, 102 videos

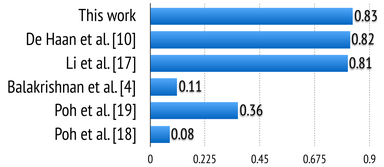


Emotion elicitation
Multiple movements and expressions
Rich temporal heart rate dynamics

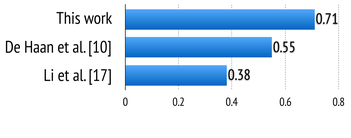
Results

Average heart rate estimation

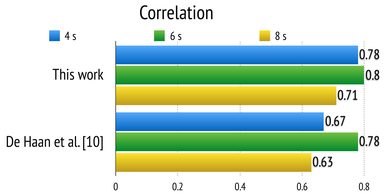
Mahnob-HCI

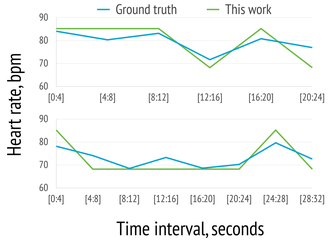


MMSE-HR

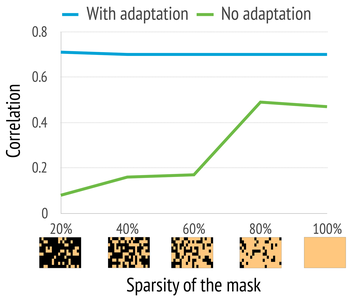
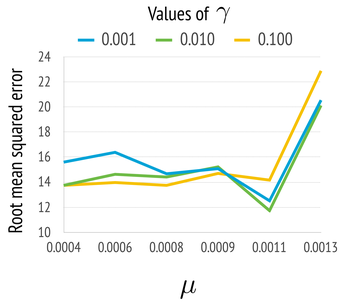


Short-term heart rate estimation





Effect of self-adaptation



Hyper parameters selection