

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



UNIVERSITY
OF TRENTO

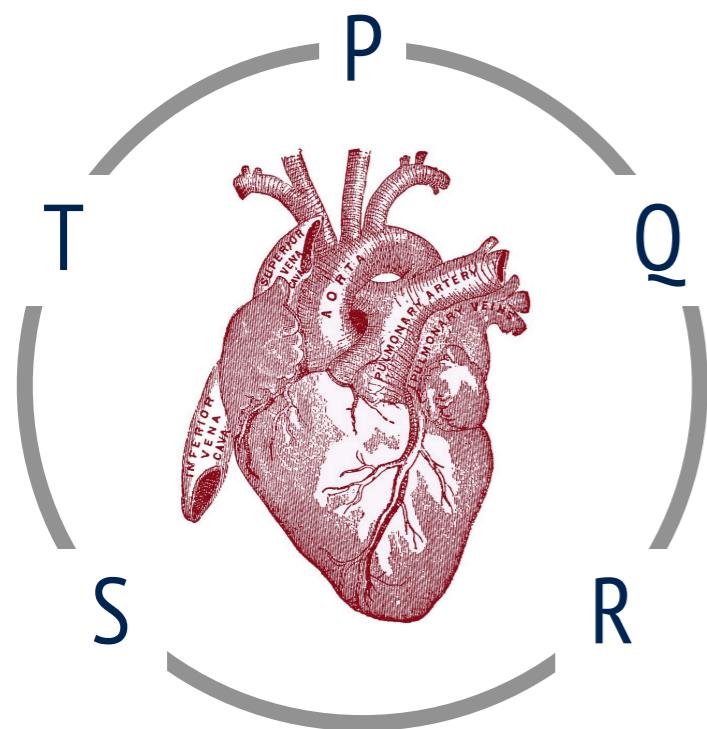
BINGHAMTON
UNIVERSITY
STATE UNIVERSITY OF NEW YORK



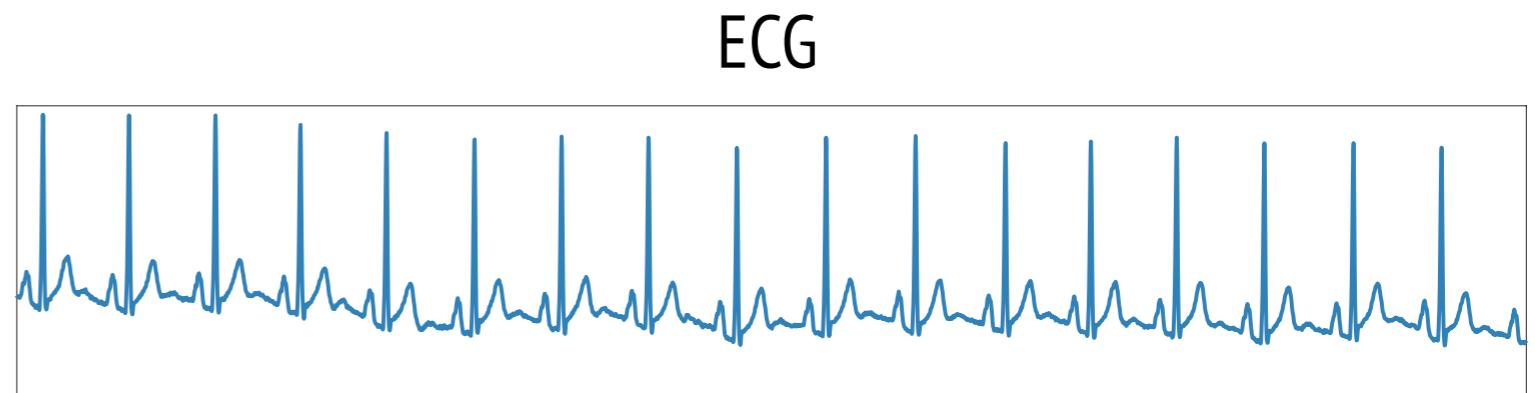
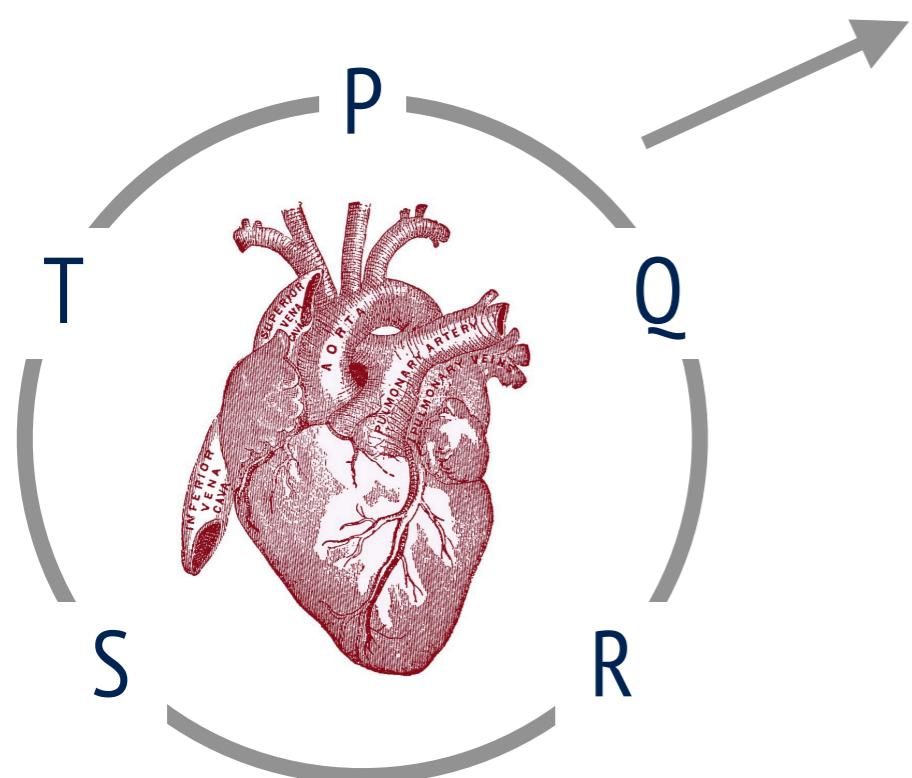
Carnegie
Mellon
University

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

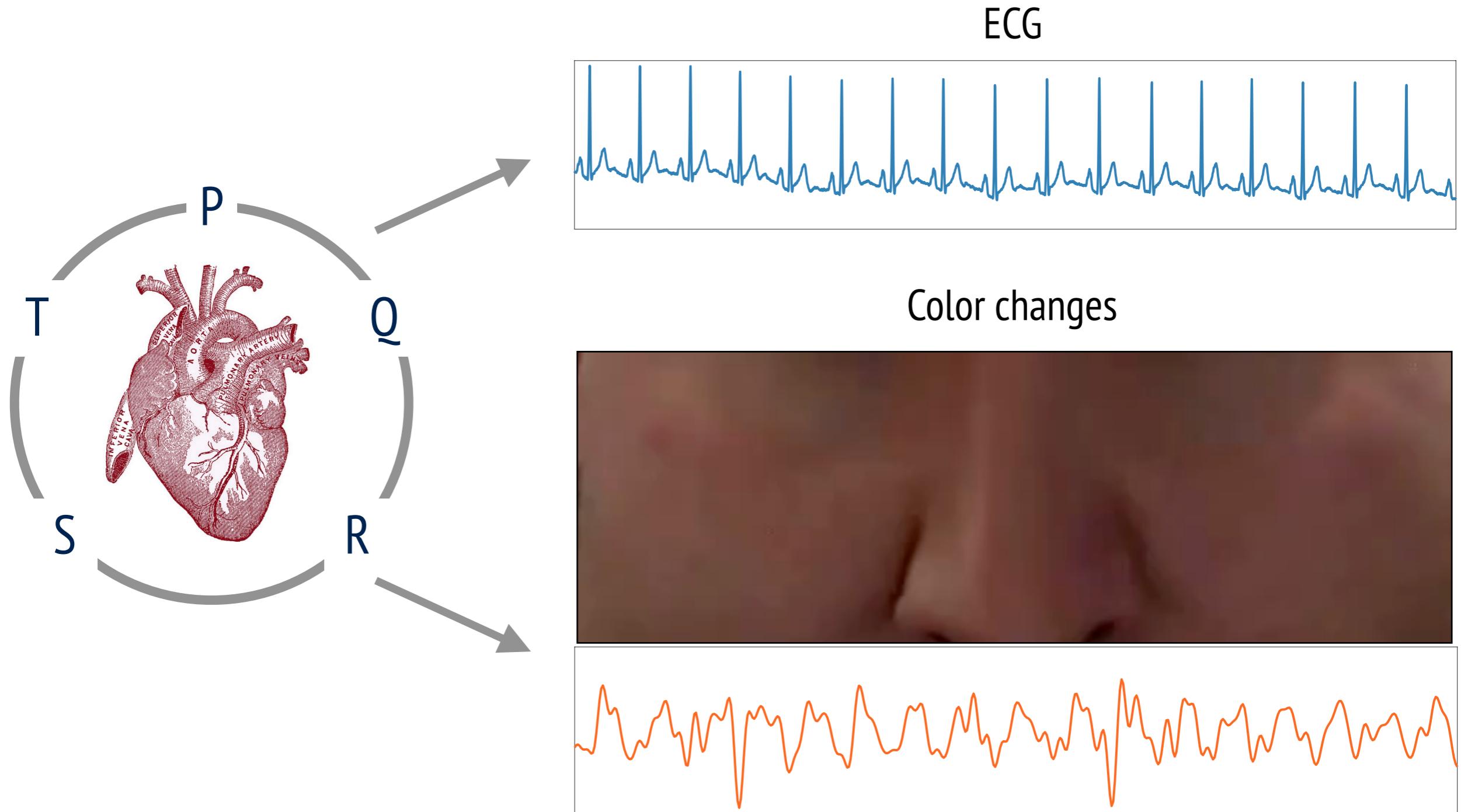
Introduction



Introduction

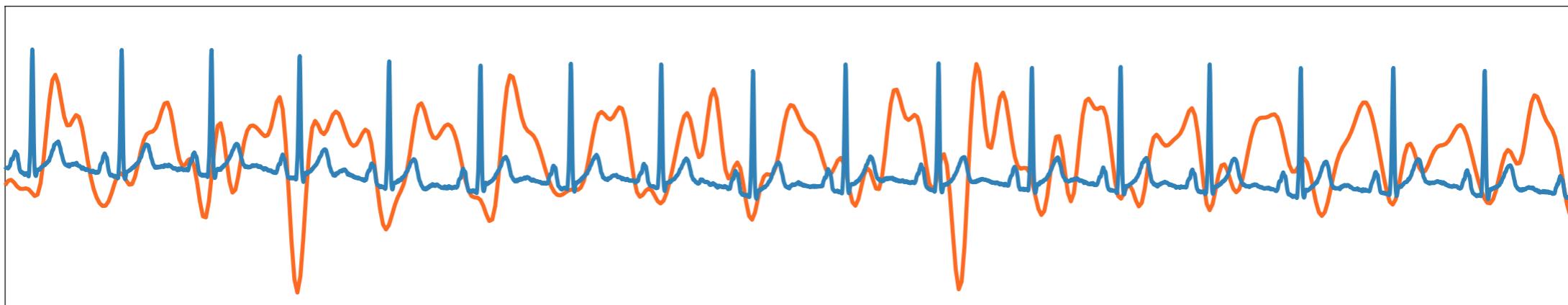


Introduction



[27] Wu et al. Eulerian video magnification for revealing subtle changes in the world. SIGGRAPH'2012

Introduction



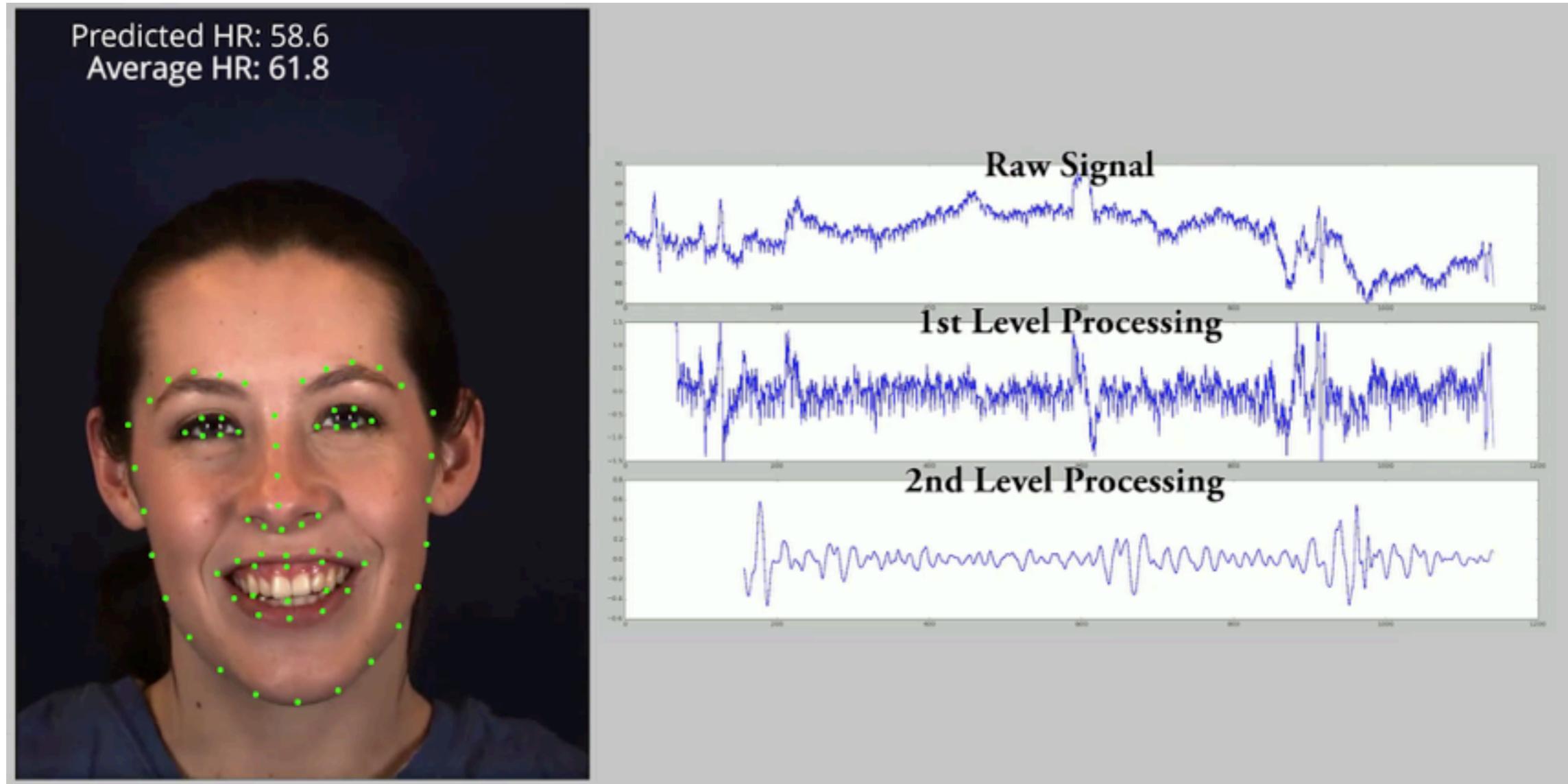
Blind source separation [18,19]

Subtle motions analysis [4]

Least-squares filtering [17]

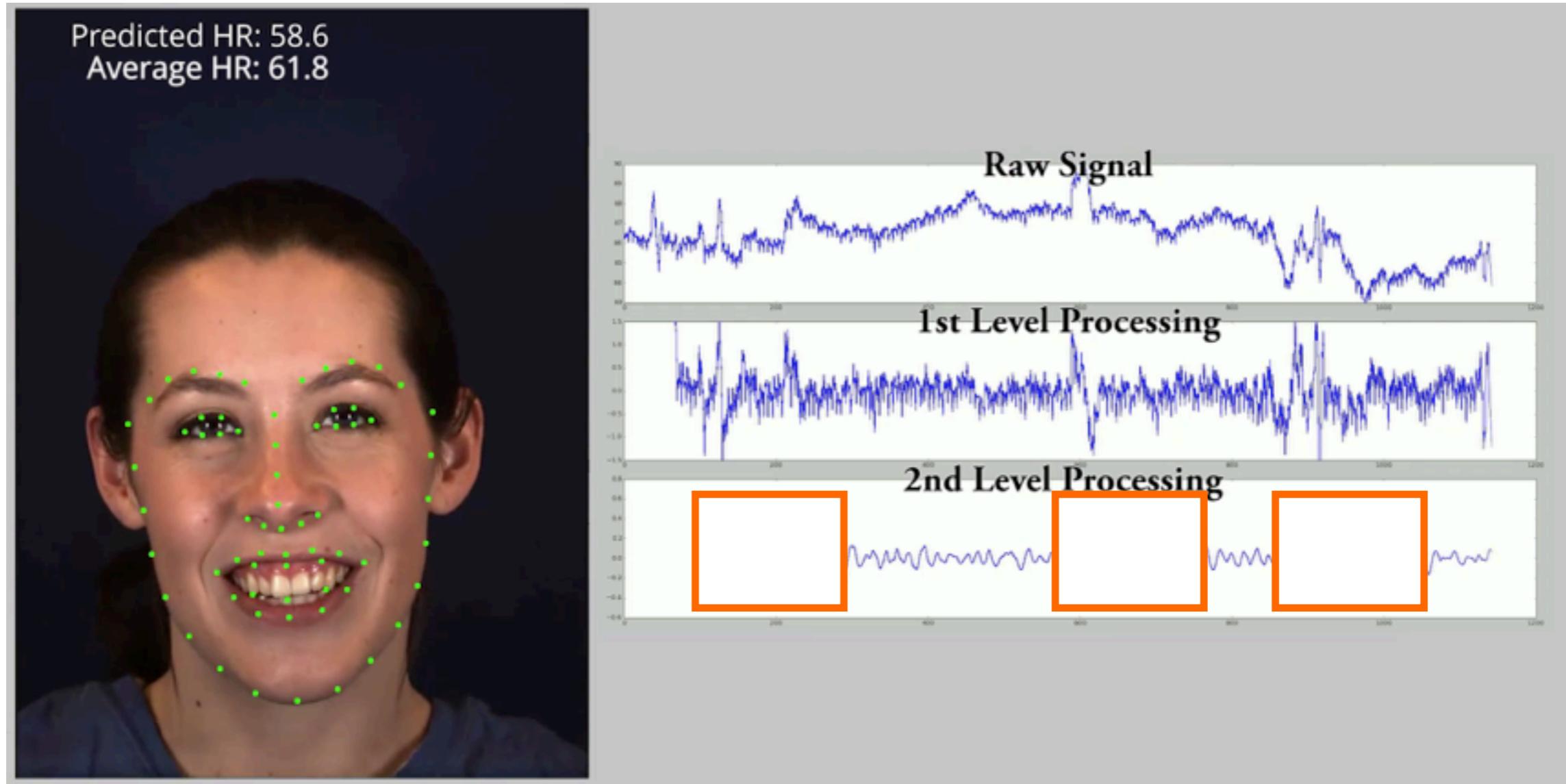
Chrominance features analysis [10, 25]

Challenging scenarios



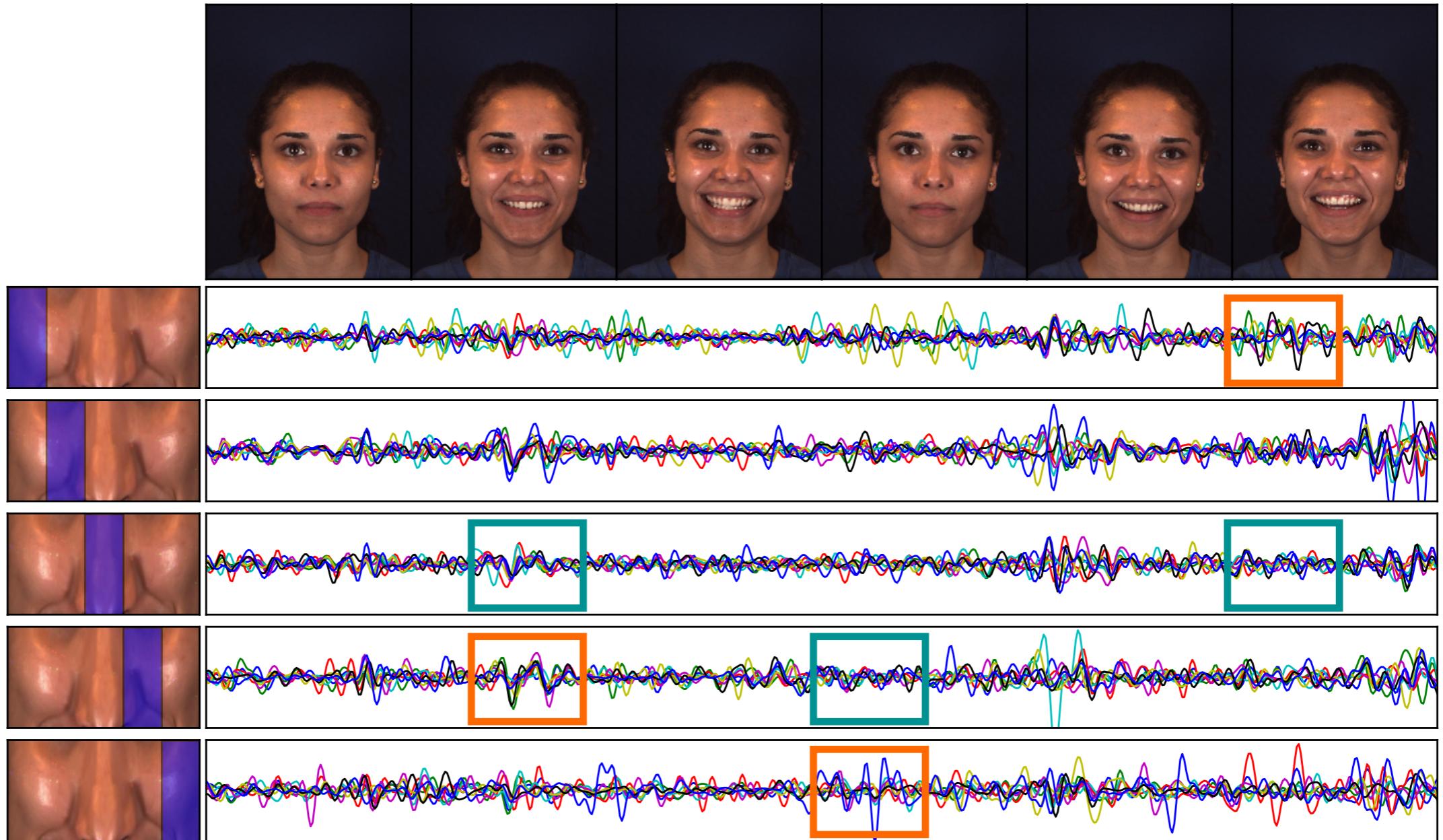
[17] Li et al. Remote Heart Rate Measurement From Face Videos Under Realistic Situations. CVPR'2014

Challenging scenarios



[17] Li et al. Remote Heart Rate Measurement From Face Videos Under Realistic Situations. CVPR'2014

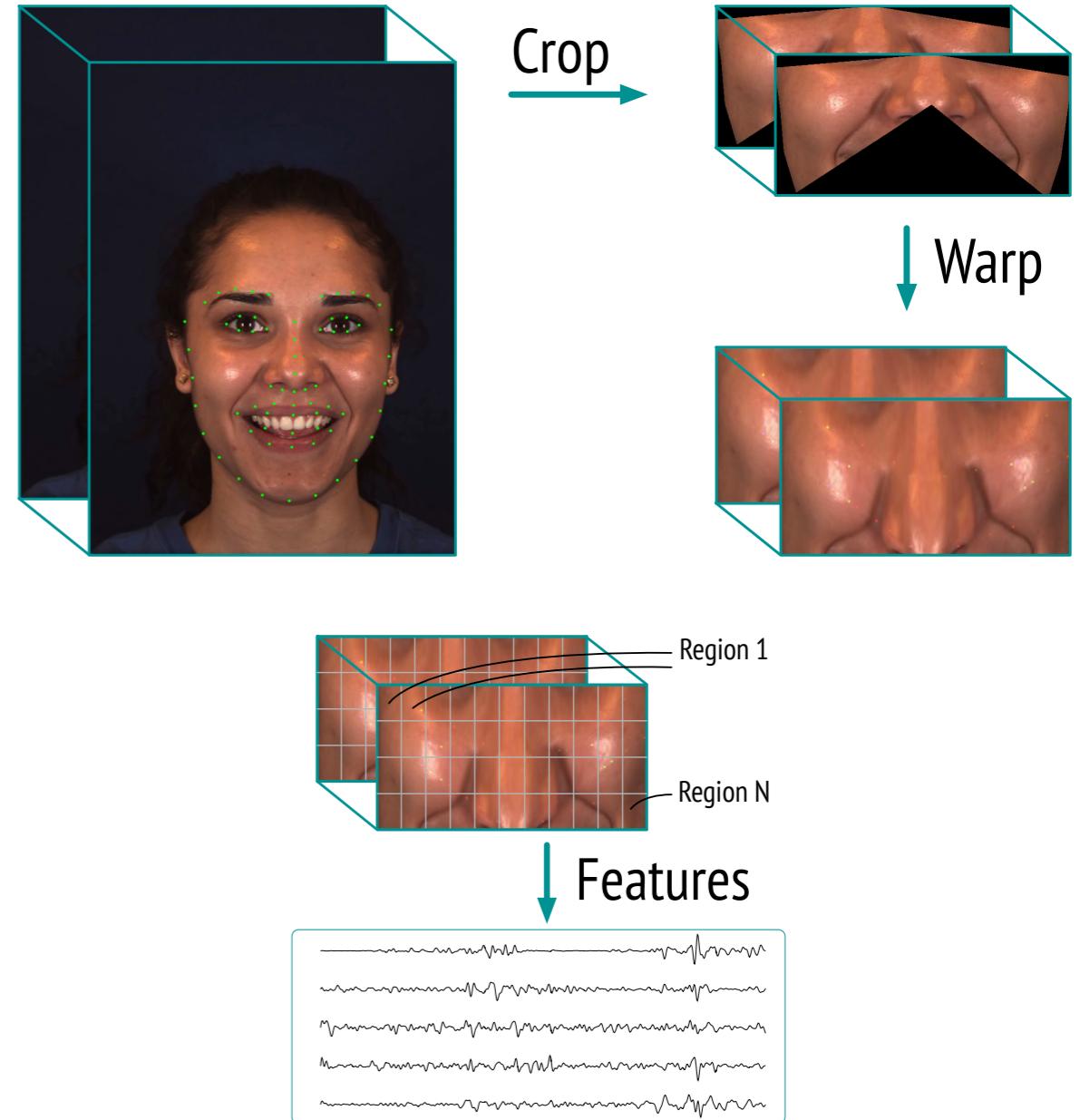
Everything is in the face



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

Heart rate features

- Track 66 points [29]
- Crop face
- Apply piece-wise linear warping
- Compute normalized chrominance features [10]

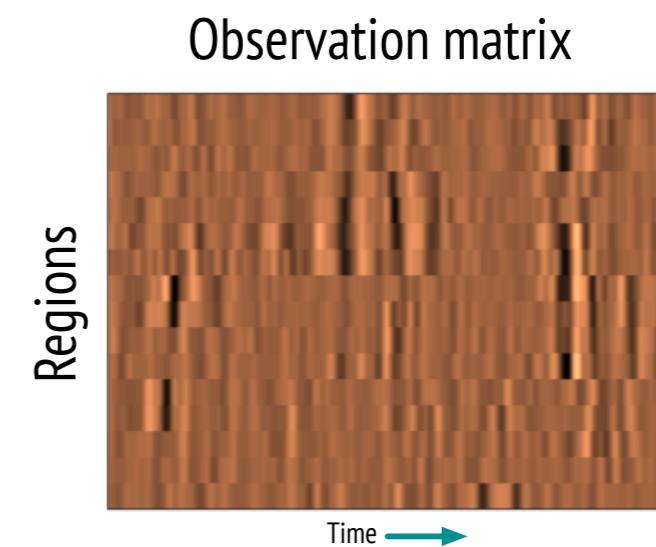
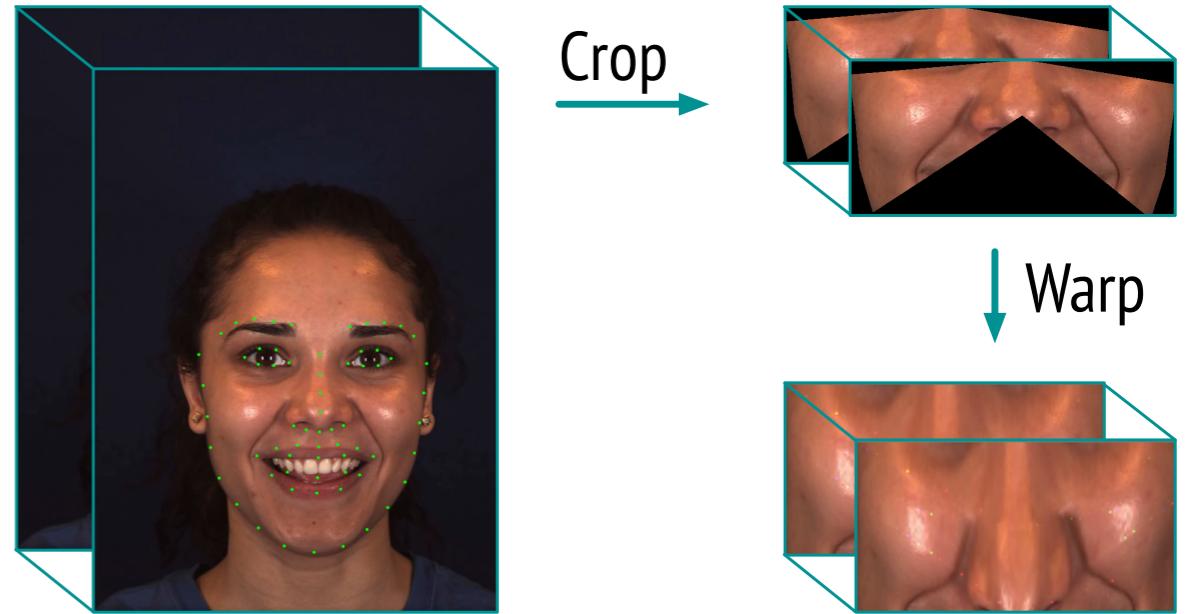
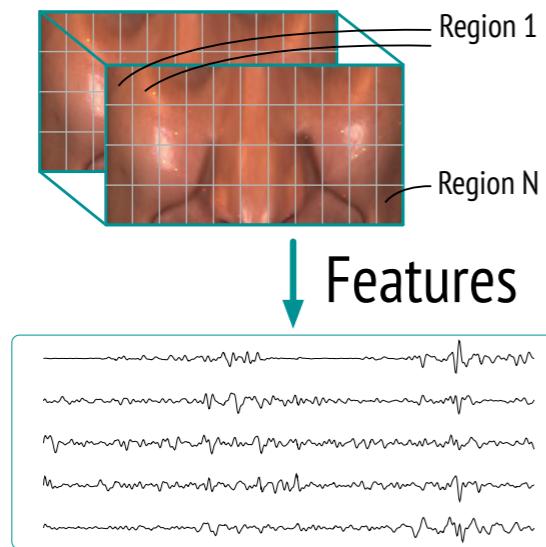


[29] Xiong et al. Supervised descent method and its applications to face alignment, CVPR'2013

[10] De Haan et al. Robust pulse rate from chrominance-based rPPG. Transactions on Affective Computing'2013

Heart rate features

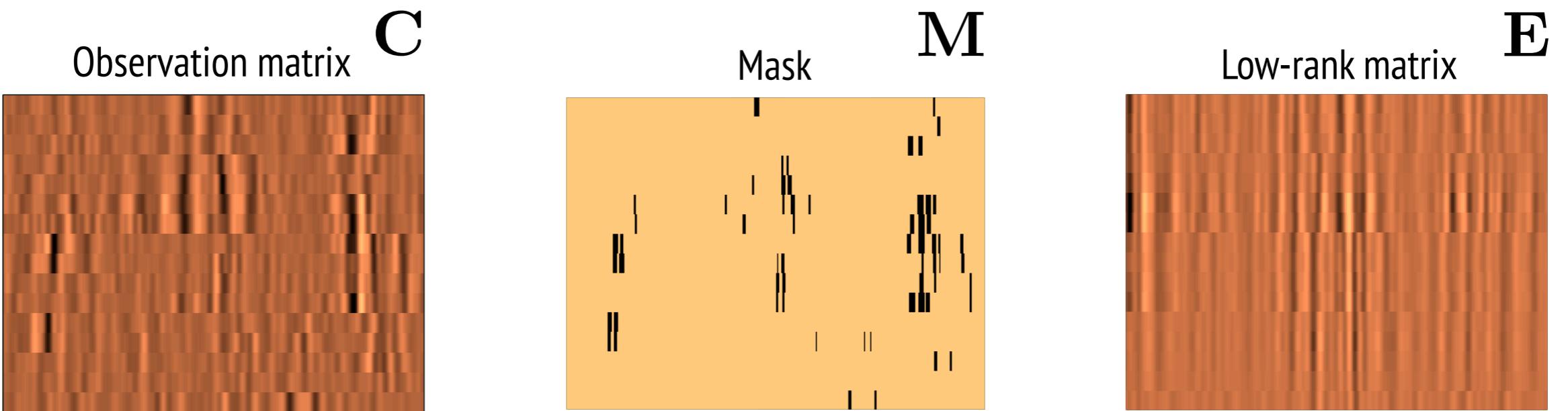
- Track 66 points [29]
- Crop face
- Apply piece-wise linear warping



- [29] Xiong et al. Supervised descent method and its applications to face alignment, CVPR'2013
[10] De Haan et al. Robust pulse rate from chrominance-based rPPG. Transactions on Affective Computing'2013

Formulation

Classical matrix completion: **given a mask** estimate a low-rank matrix



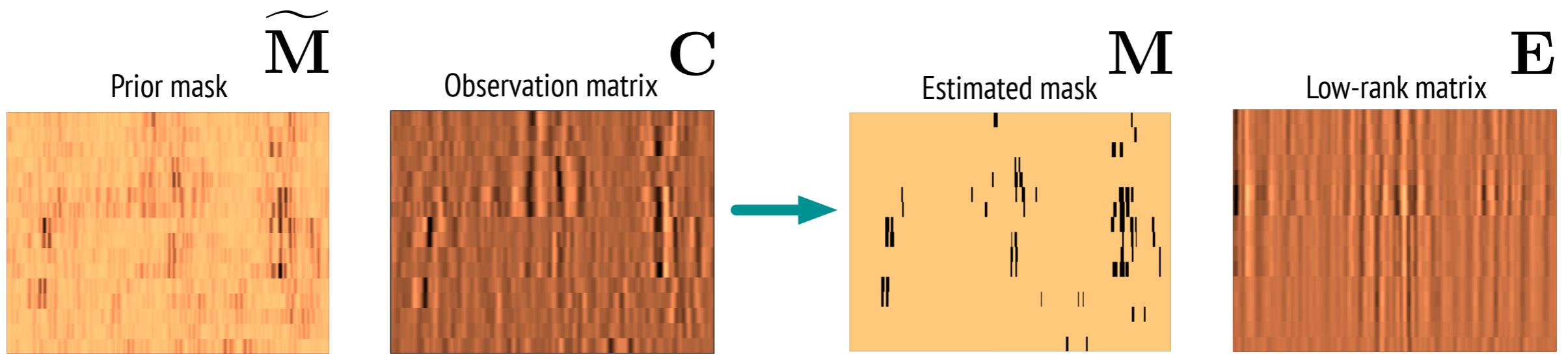
Solve

$$\min_{\mathbf{E}} \nu \|\mathbf{E}\|_* + \|\mathbf{M} \circ (\mathbf{E} - \mathbf{C})\|_{\mathcal{F}}^2$$

The mask is fixed

Self adaptive matrix completion

Goal: jointly estimate the low rank matrix and the mask



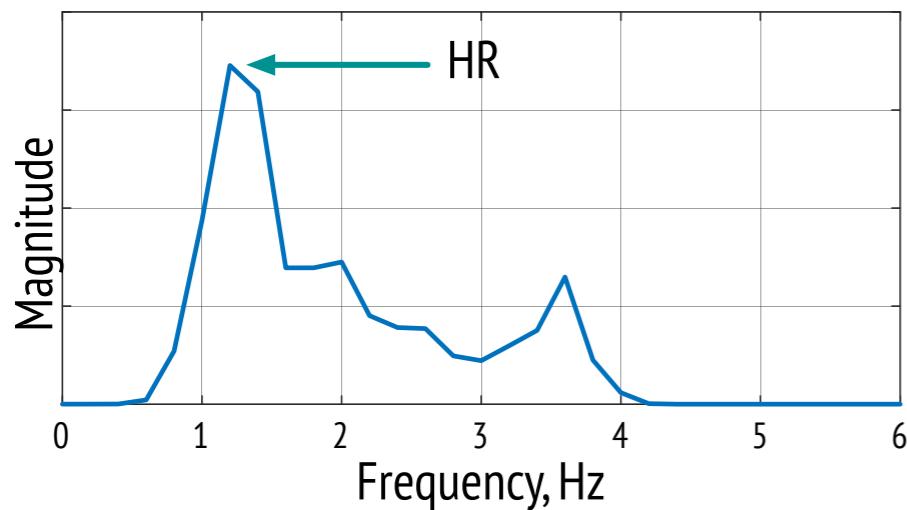
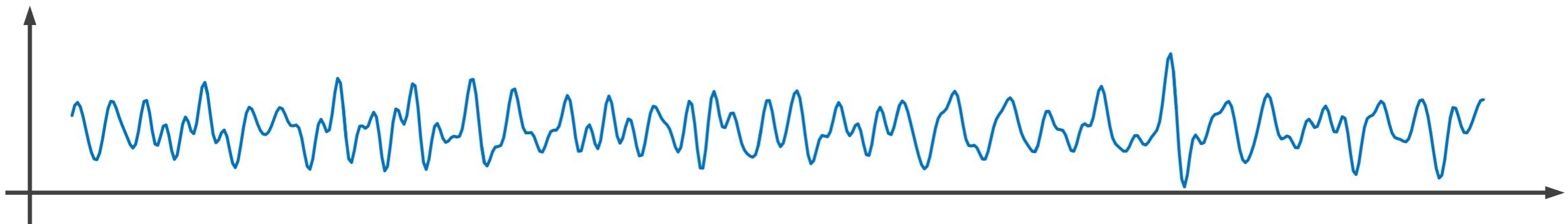
Jointly solve

$$\begin{aligned} \min_{\mathbf{E}, \mathbf{M}} \quad & \nu \|\mathbf{E}\|_* + \|\mathbf{M} \circ (\mathbf{E} - \mathbf{C})\|_{\mathcal{F}}^2 + \\ & \boxed{\gamma \text{Tr}(\mathbf{E} \mathbf{L} \mathbf{E}^\top)} - \boxed{\beta \|\mathbf{M}\|_1} + \mu \|\mathbf{M} - \tilde{\mathbf{M}}\|_{\mathcal{F}}^2, \end{aligned}$$

Temporal smoothing **Anti-sparsity**

Heart rate estimation

Compute signal associated with the largest singular value



68
beats
per
minute

[19] P. D. Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. Transactions on Audio and Electroacoustics'1967

MAHNOB-HCI dataset



27 subjects
20 videos per subject

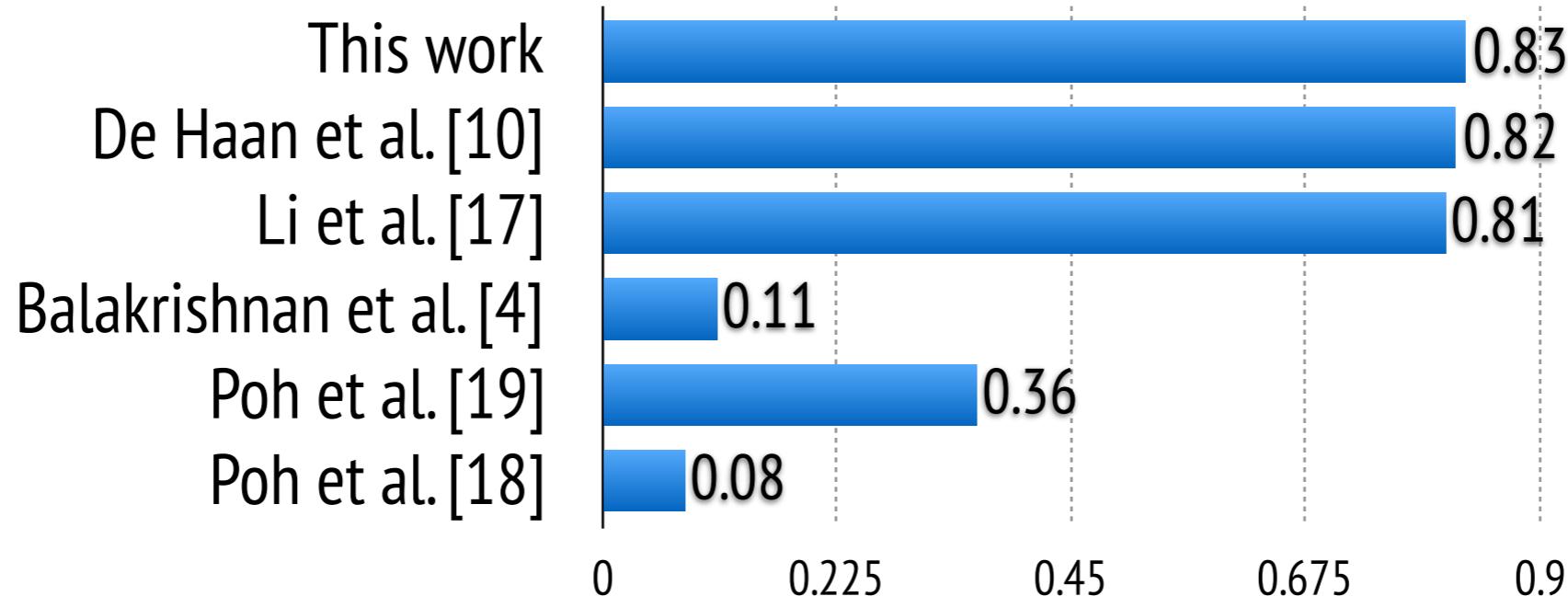
[20] Soleymani et al. A Multimodal Database for Affect Recognition and Implicit Tagging. Transactions on Affective Computing'2012

MAHNOB-HCI dataset



27 subjects
20 videos per subject

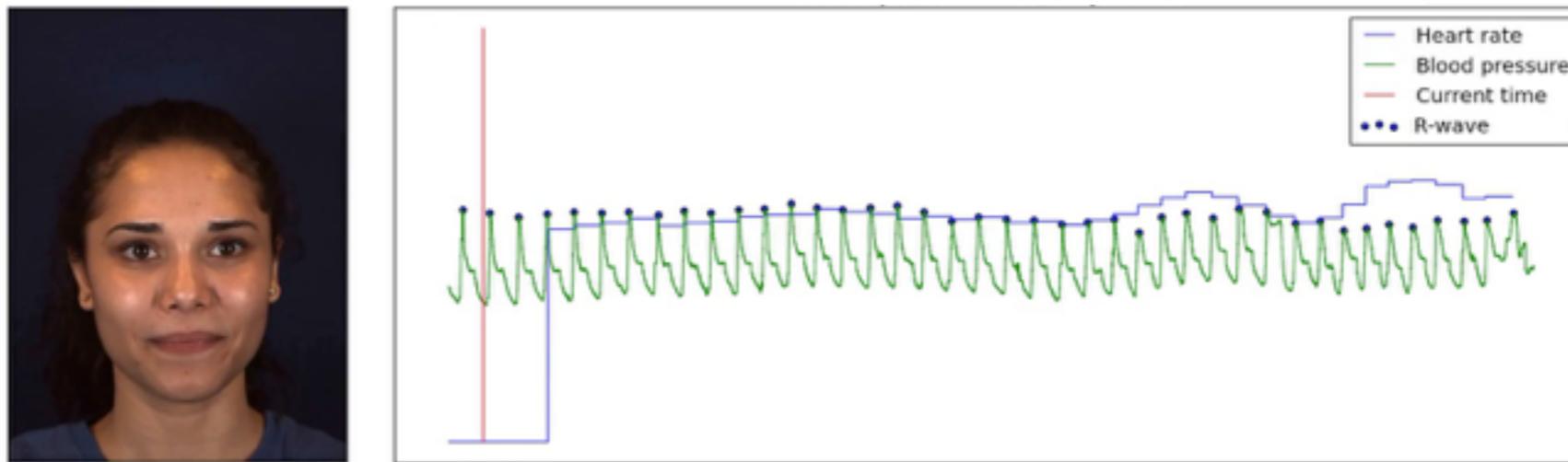
Correlation



[20] Soleymani et al. A Multimodal Database for Affect Recognition and Implicit Tagging. Transactions on Affective Computing'2012

MMSE-HR dataset

Time for fear. They throw darts in her face



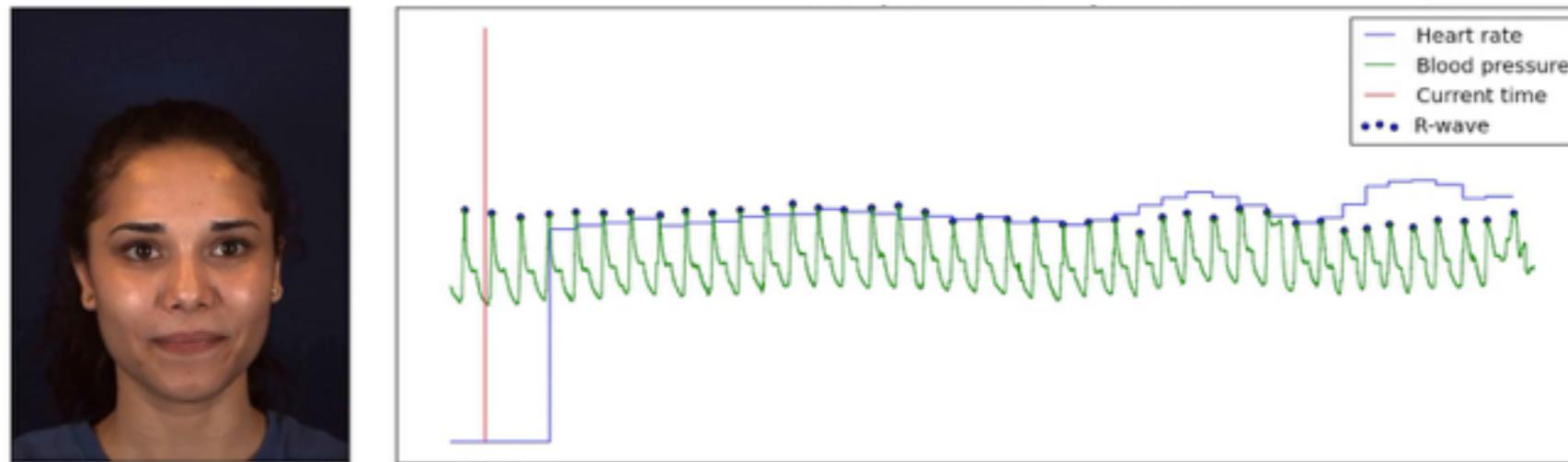
- 40 subjects / 102 videos
- RGB physiological data
- Elicited emotions

MMSE at CVPR
June 29
Poster #44, session P3-1

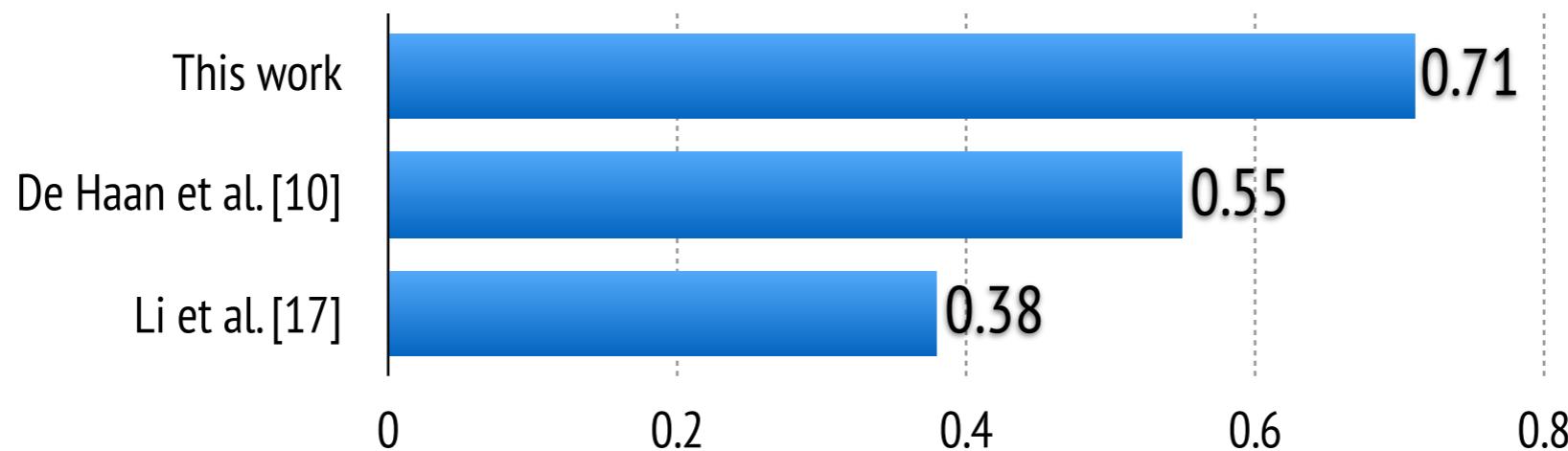
[31] Zhang et. al. Multimodal Spontaneous Emotion Corpus for Human Behavior Analysis, CVPR'2016
The dataset supported in part by U.S. National Science Foundation grants CNS-1205664 and CNS-1205195

MMSE-HR dataset

Time for fear. They throw darts in her face

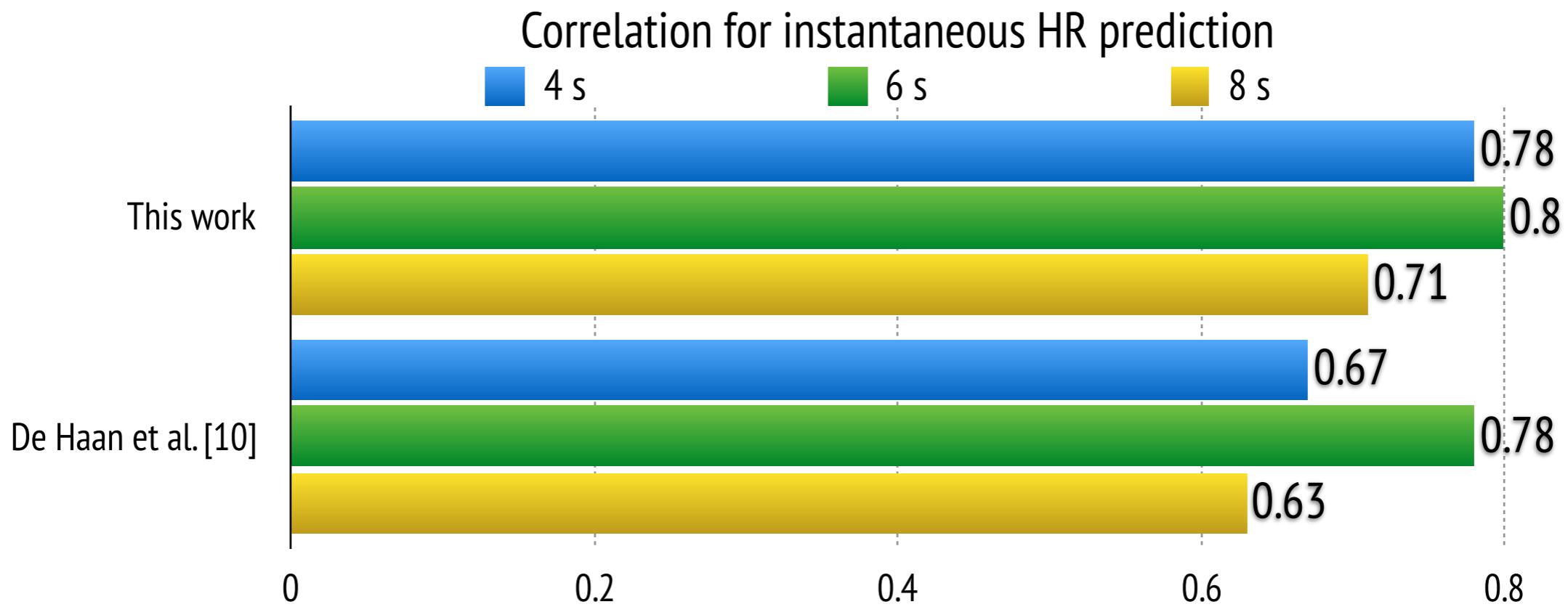
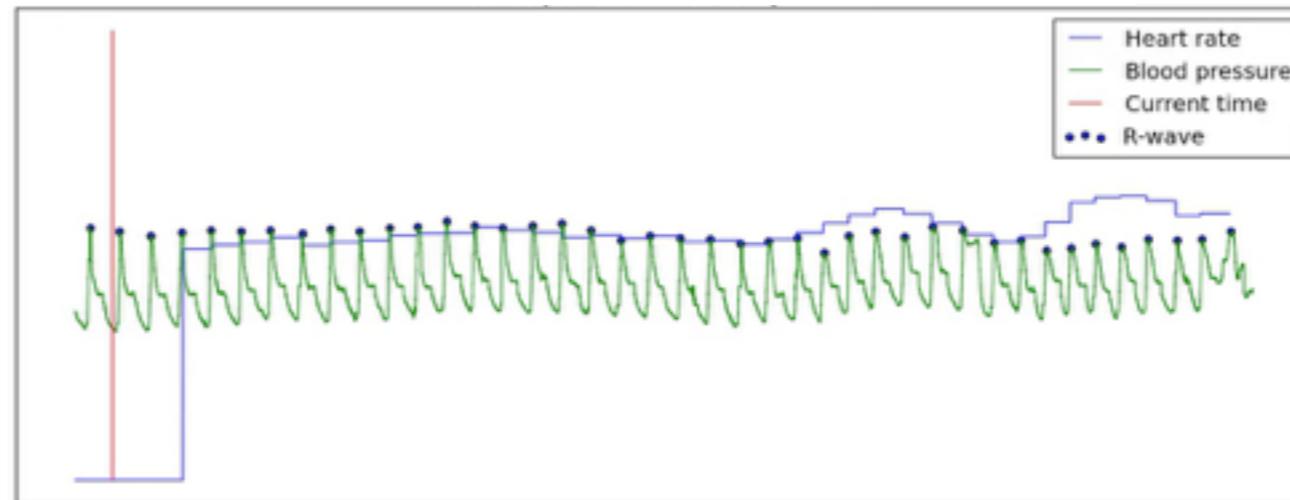
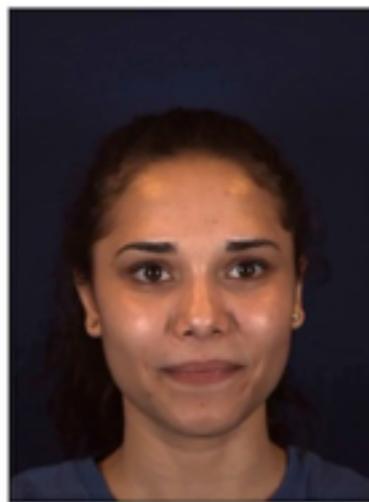


Correlation for average HR prediction

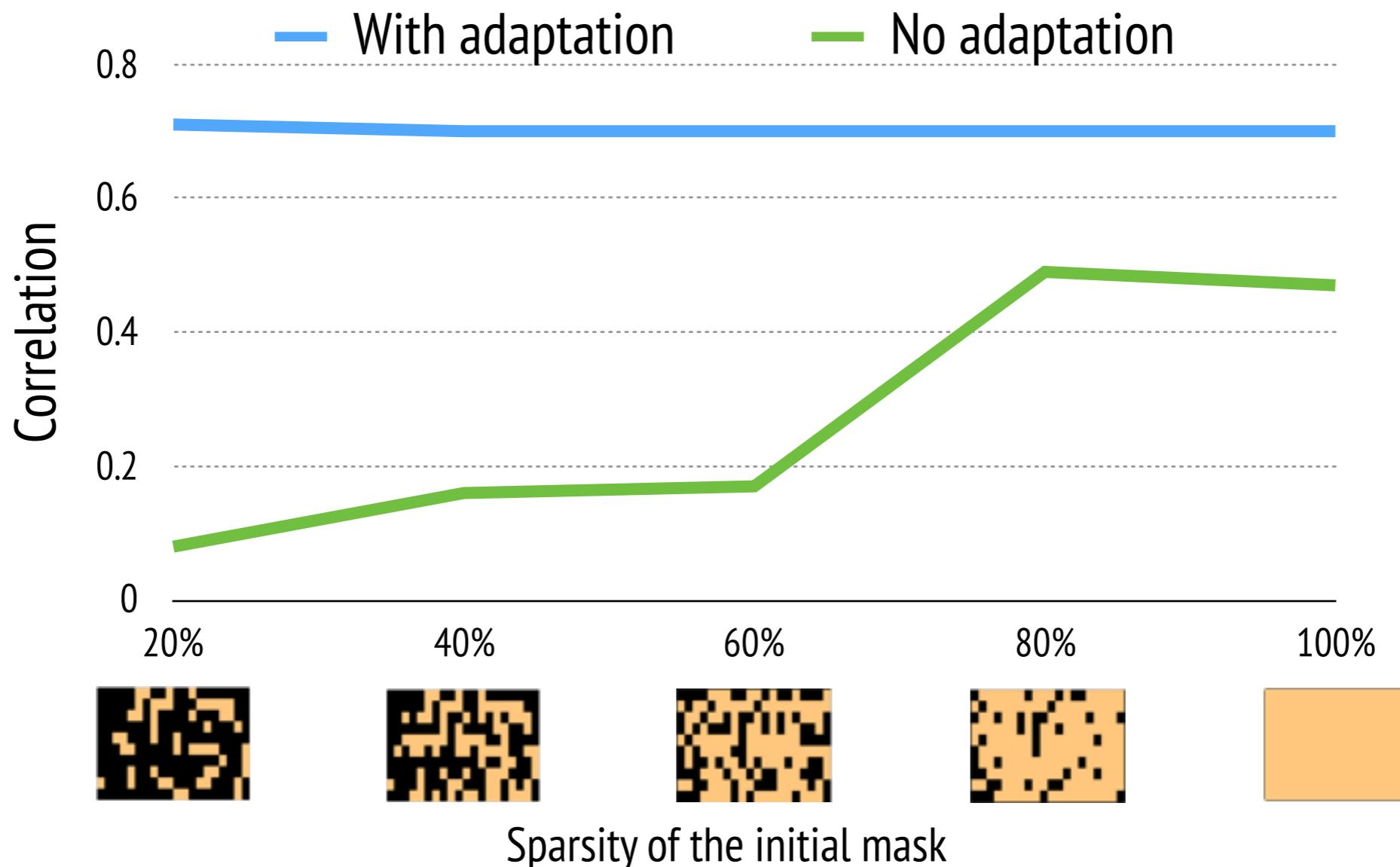


MMSE-HR dataset

Time for fear. They throw darts in her face

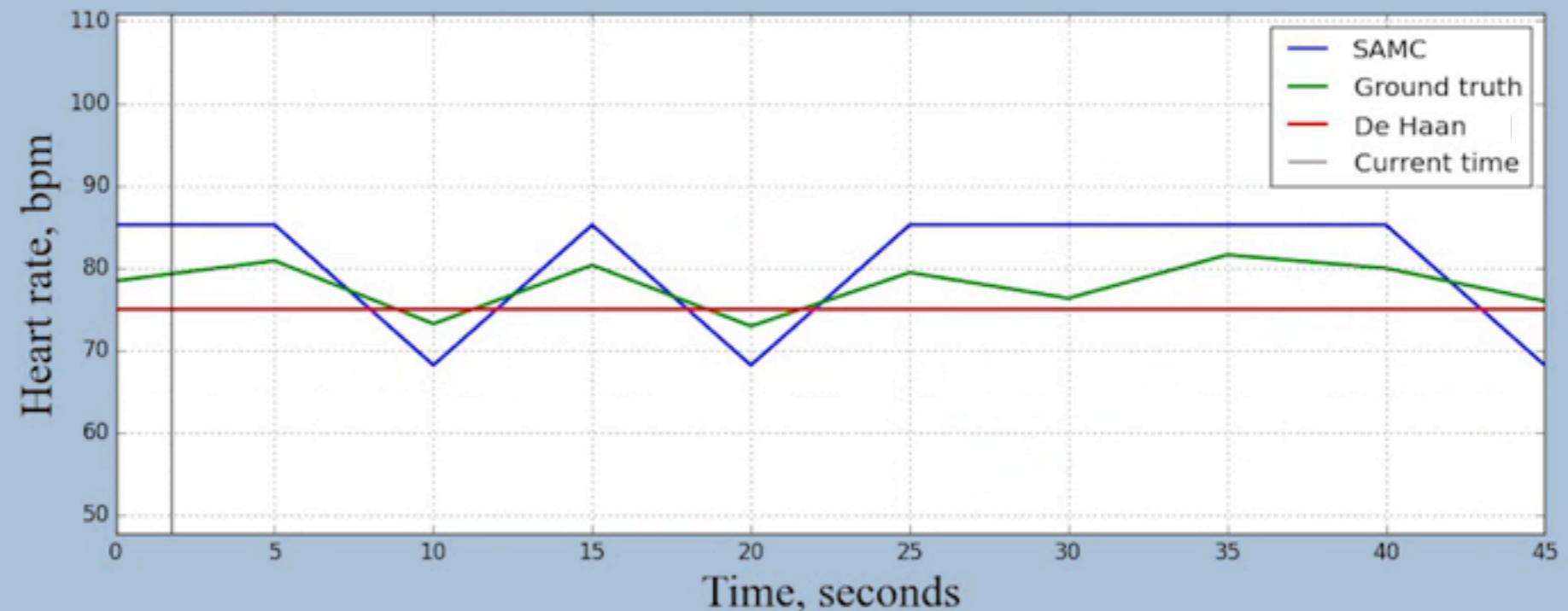


Effect of self adaptation



Short-time HR estimation

Subject M002 T11. Pain Sequence
Time window 4 seconds. Speed 1.2



Real-life demo



Pre-brexit Cameron

Take home message:

- Face contains the necessary information
- Discarding noisy face regions helps
- SAMC shows remarkable performance on HR estimation

Future work:

- Multimodal HR estimation
- Other physiological / emotional signals