Comparison of Face Segmentation Methods for Non-contact Video-based Pulse Rate Measurement

Ronny Stricker
Ilmenau University of Technology
Neuroinformatics and Cognitive Robotics Lab
Ronny.Stricker@tu-ilmenau.de
www.tu-ilmenau.de/neurob

S. Müller and H.-M. Gross

Ilmenau Univ. of Technology
Germany

Outline

- Project Goals
- Principle of Photoplethysmography
- Processing Pipeline
- ROI Extraction Methods
- Pulse Database
- Results and Outlook

Project Goals

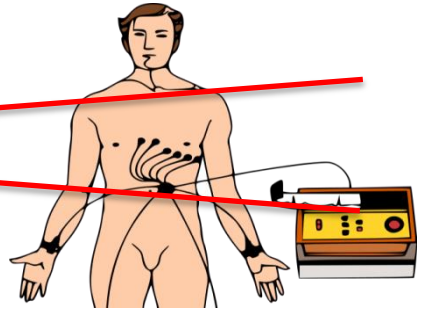
- Prevention and **assistance for elderly** people living alone in their **home environment**
- In the scope of different research projects (e.g. SERROGA or Roreas)

- The robots main purposes are to
  - serve as **“communication assistant”**
  - give **“motion motivation”**
  - provide **”reminder service”**

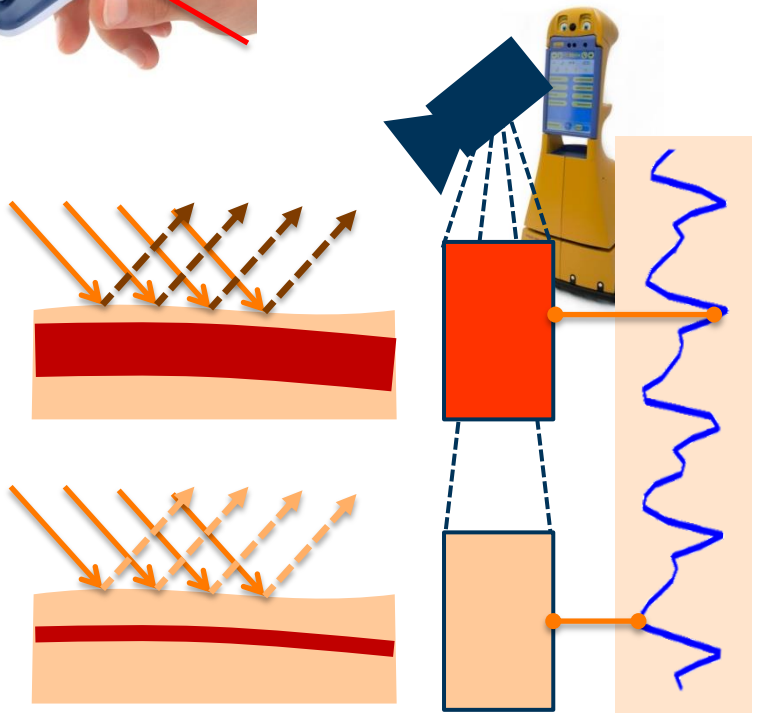
- Furthermore it should provide **basic health monitoring**
- **nonintrusive pulse measurement** can be very valuable

Measuring the human pulse

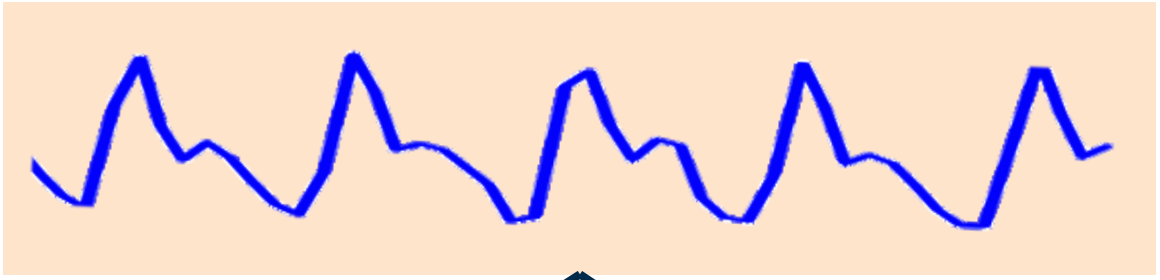
- ~~Pulse rate can be measured reliably using **electrocardiography** (ECG) methods -> Gold standard but very intrusive~~

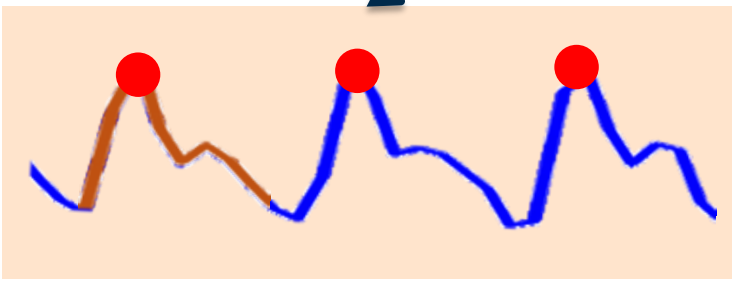


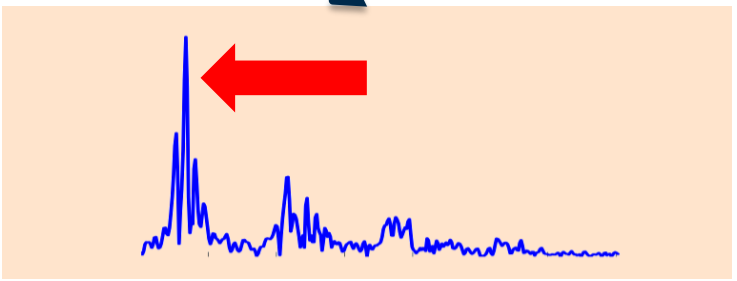
- **Photoplethysmography** is an alternative method for pulse rate measurements
  - History
    - ~~Pulse oximeter (Goldberger et al. 1987)~~
    - ~~Remote with Active red, infrared (2005)~~
    - Remote with ambient light (2008)
  - Light is absorbed stronger by blood than by the surrounding tissue
  - Reflectance of skin changes according to the changes of blood volume in the microvascular bed of tissue



Measuring the human pulse







Temporal Domain

Frequency Domain

Processing Pipeline

ROI Extraction

Face Detection

Skin Color Segmentation

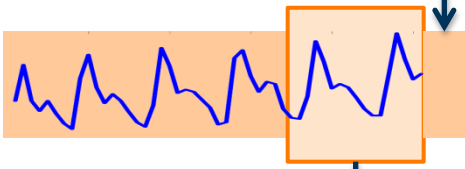
⋮

Average Pixel Values

51, 83, 14

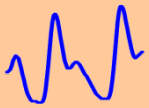
Color Space Conversion

71

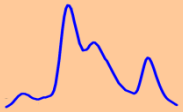


Sliding window

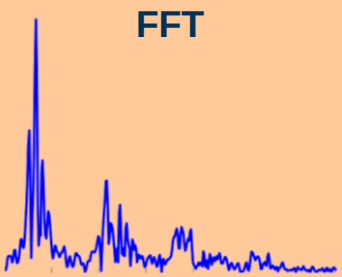
Resample



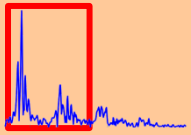
Window Function



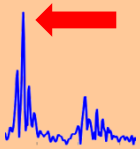
FFT



Band Pass



Arg Max

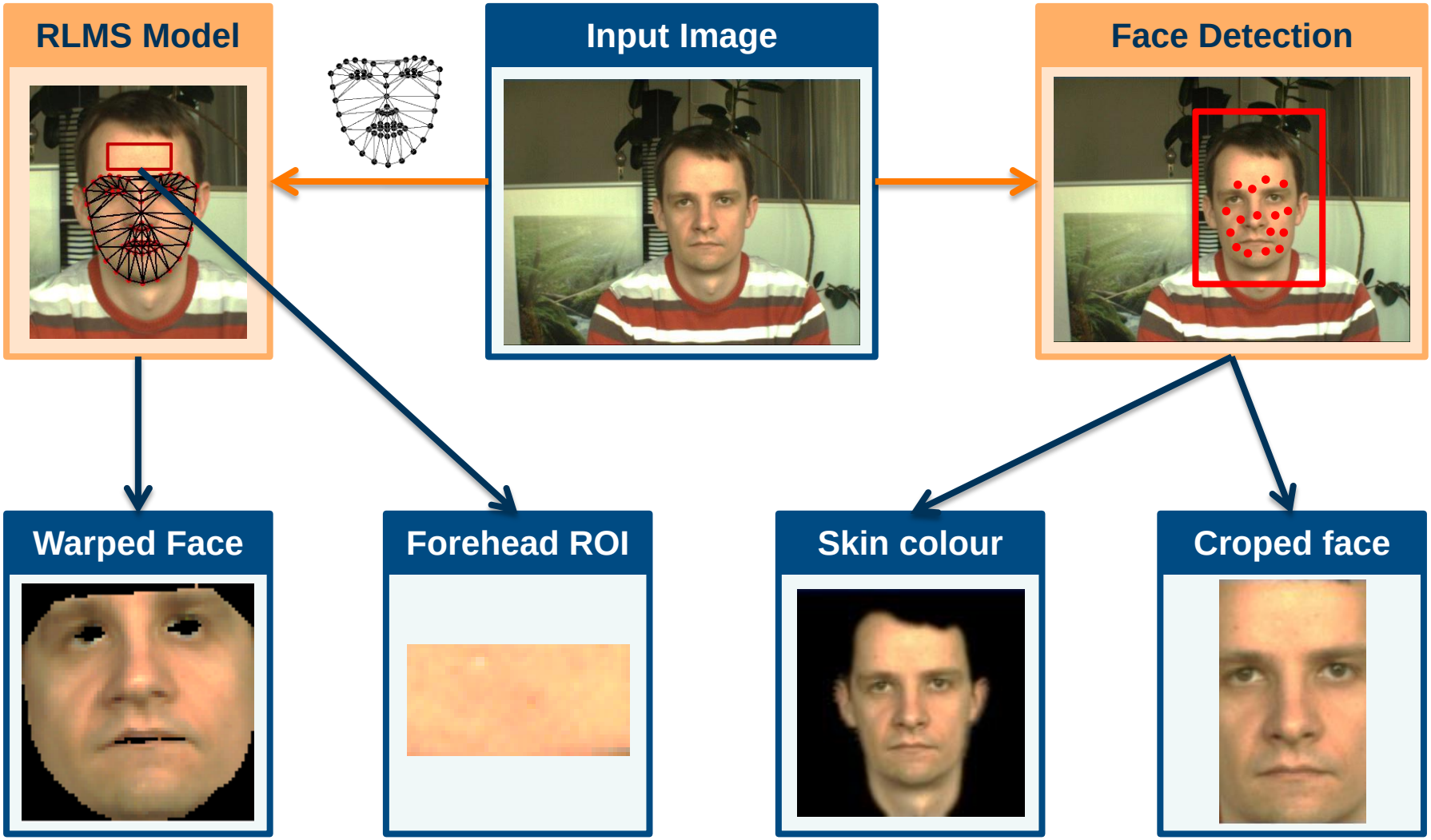


Buffer over time

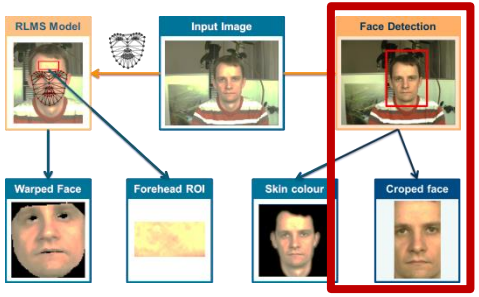
Pulse

♥ **65** BPM

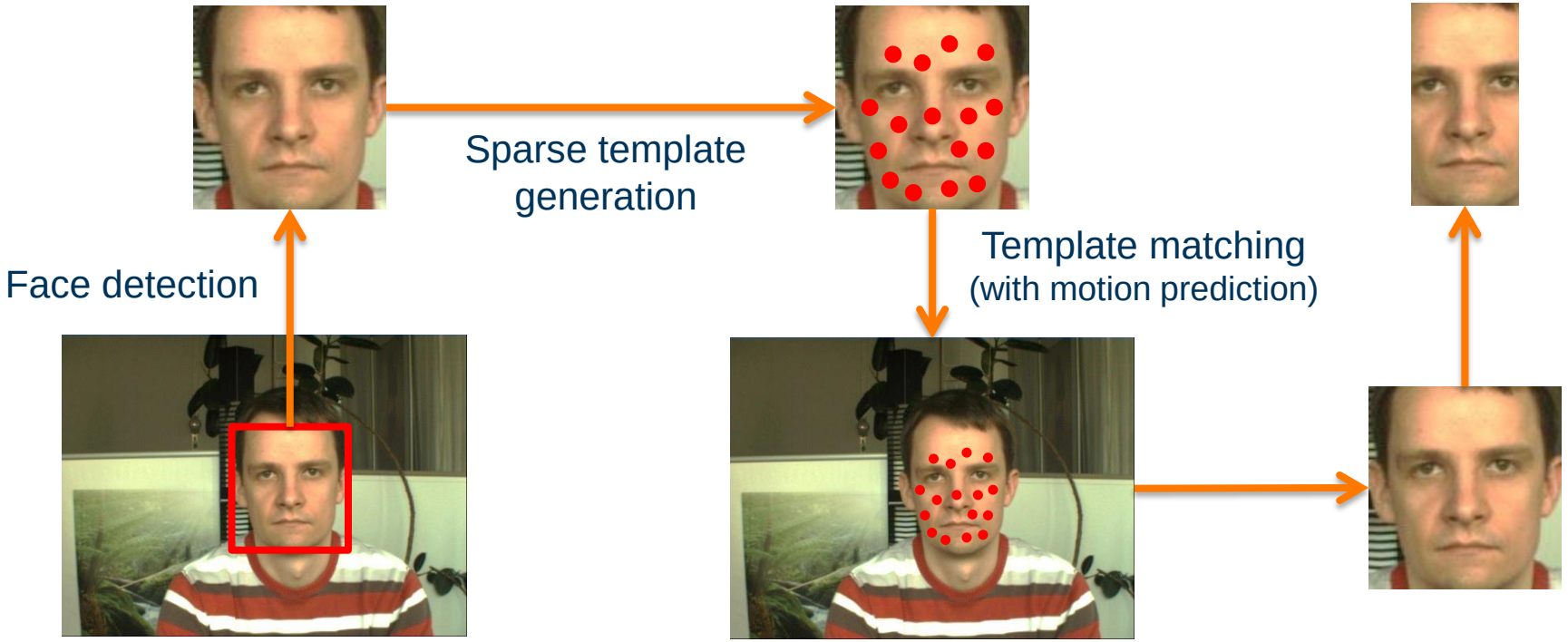
ROI Extraction - Overview



ROI Extraction – Face Detection

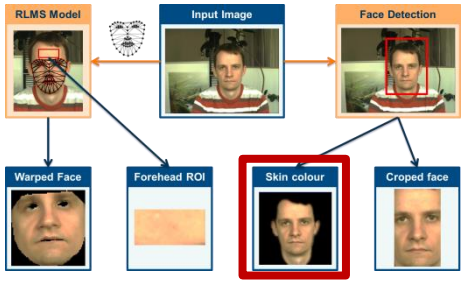


- face detection / tracking to reduce influence of background
- Subsequent face detections can disturb pulse estimates (detection ROIs are bound to discrete scales)
- Make use of face detection followed by sparse template-based feature tracker
- Can reliably track in plane object movements
- Bounding box of tracker is reduced to 60% of original width

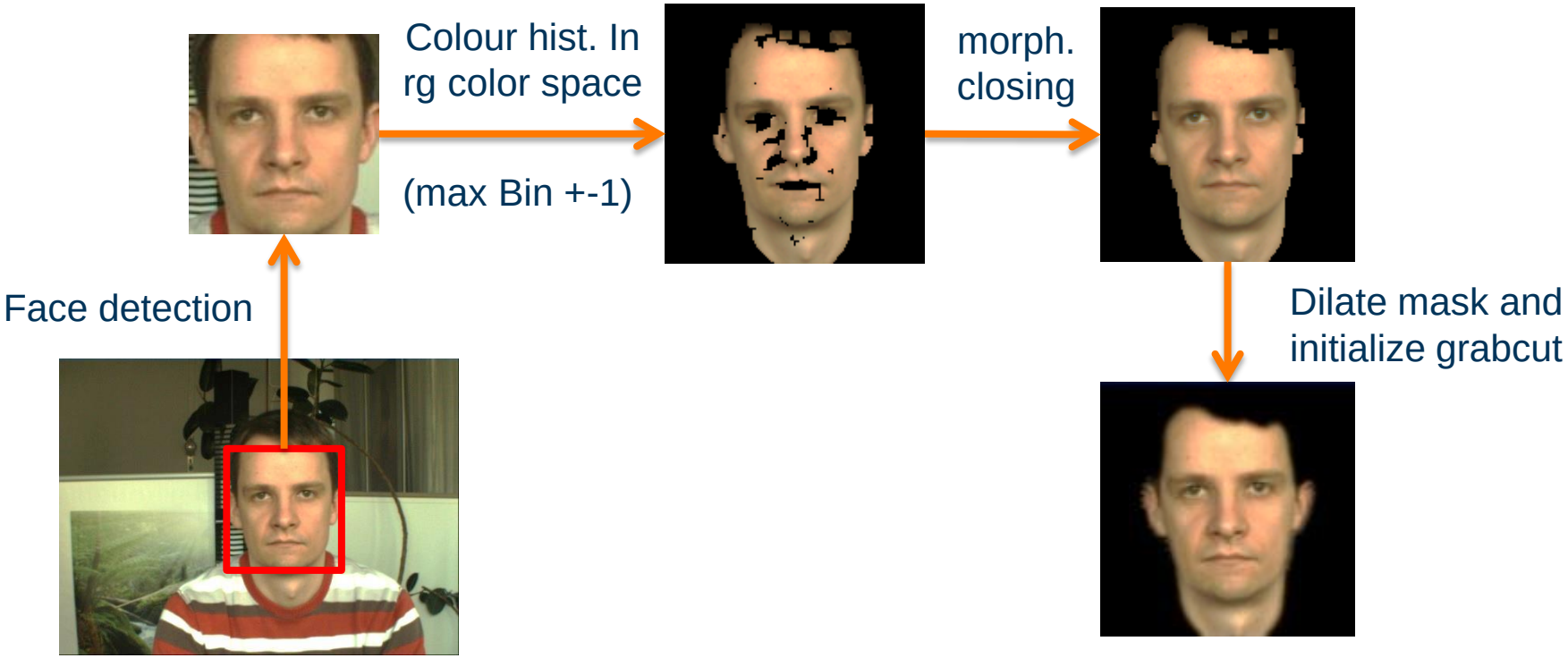


Kolarow et al., Vision-based hyper-real-time object Tracker for human-robot applications, IROS 2012

ROI Extraction – Skin Colour

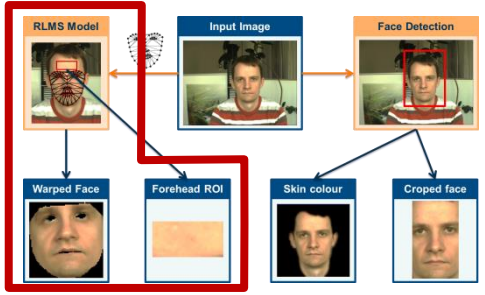


- Use grabcut (graph cut based image segmentation) to extract face skin
- Use skin colour segmentation as initialization for grabcut (foreground-background initialization)

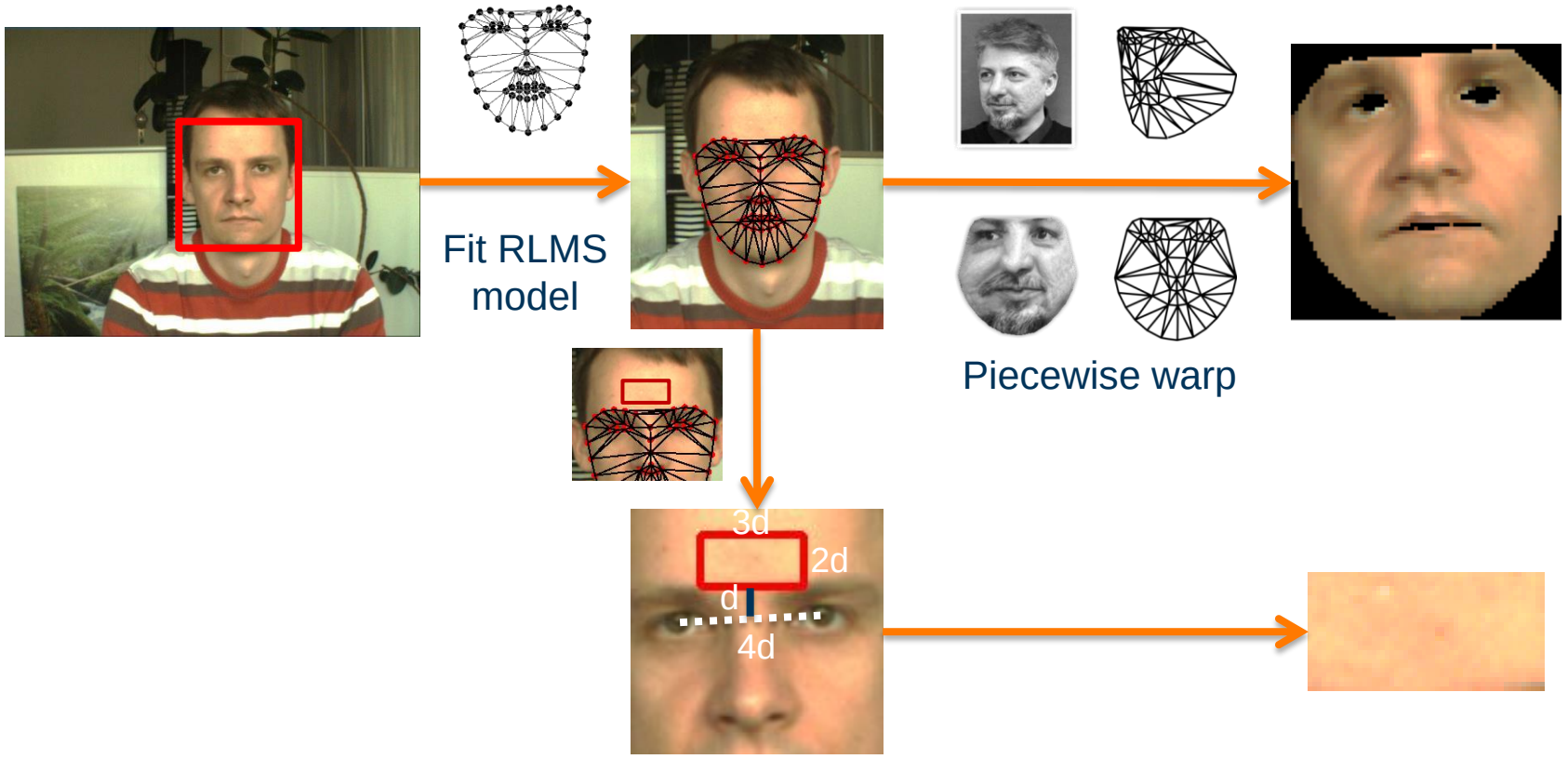


Rother et al., GrabCut: Interactive foreground extraction using iterated graph cuts, ACM Trans. Graph., 2004,

ROI Extraction – RLMS



- Use Deformable Model Fitting by Regularized Landmark Mean-Shift (RLMS) method to obtain accurate face tracking
- Use piecewise affine warp to eliminate facial movements
- RLMS is used as starting point for forehead tracking

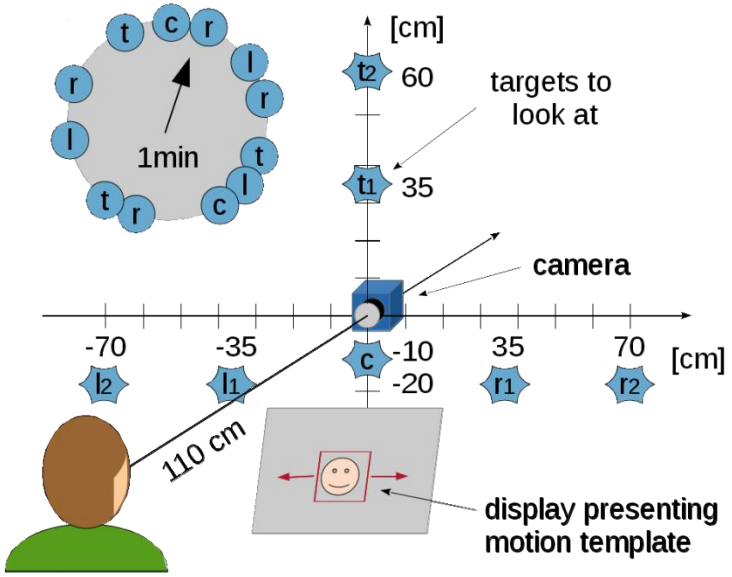


Saragih et al., Deformable model fitting by Regularized Landmark Mean-Shift, IJCV 2011

Lewandowska et al., Measuring pulse rate with a webcam – a non-contact method for evaluating cardiac activity, CCSIS 2011

Experimental Setup

- Record **database** to examine the artifacts introduced by head motion in more detail
- Head movements performed under **controlled** and **well defined** parameters
- **10 persons** (8 male, 2 female) that were recorded in **6 different setups** of 1 minute each

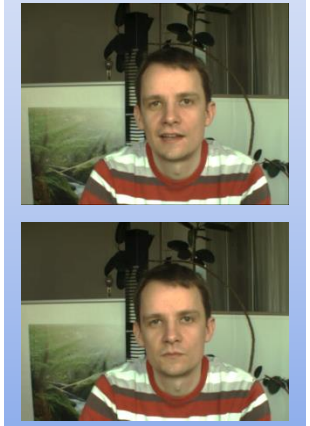


Steady

Slow/Fast translation

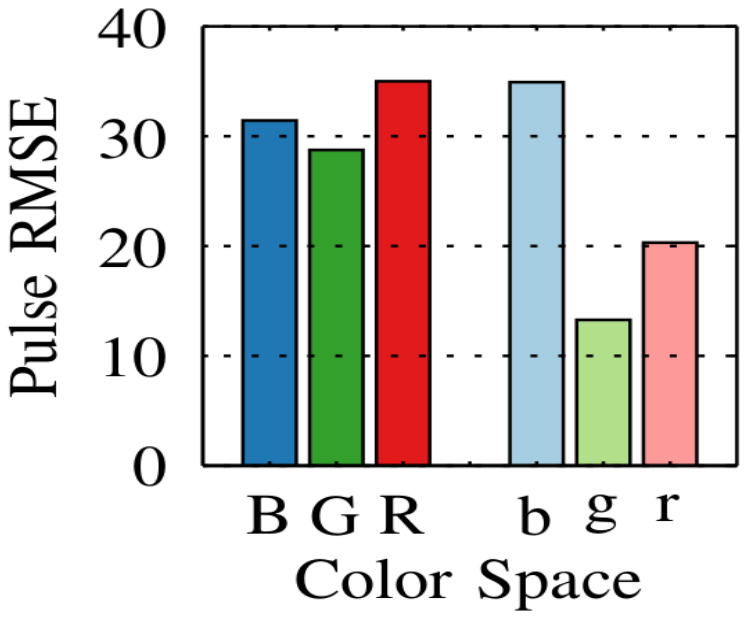
Small/Medium rotation

Talking



Colour Space

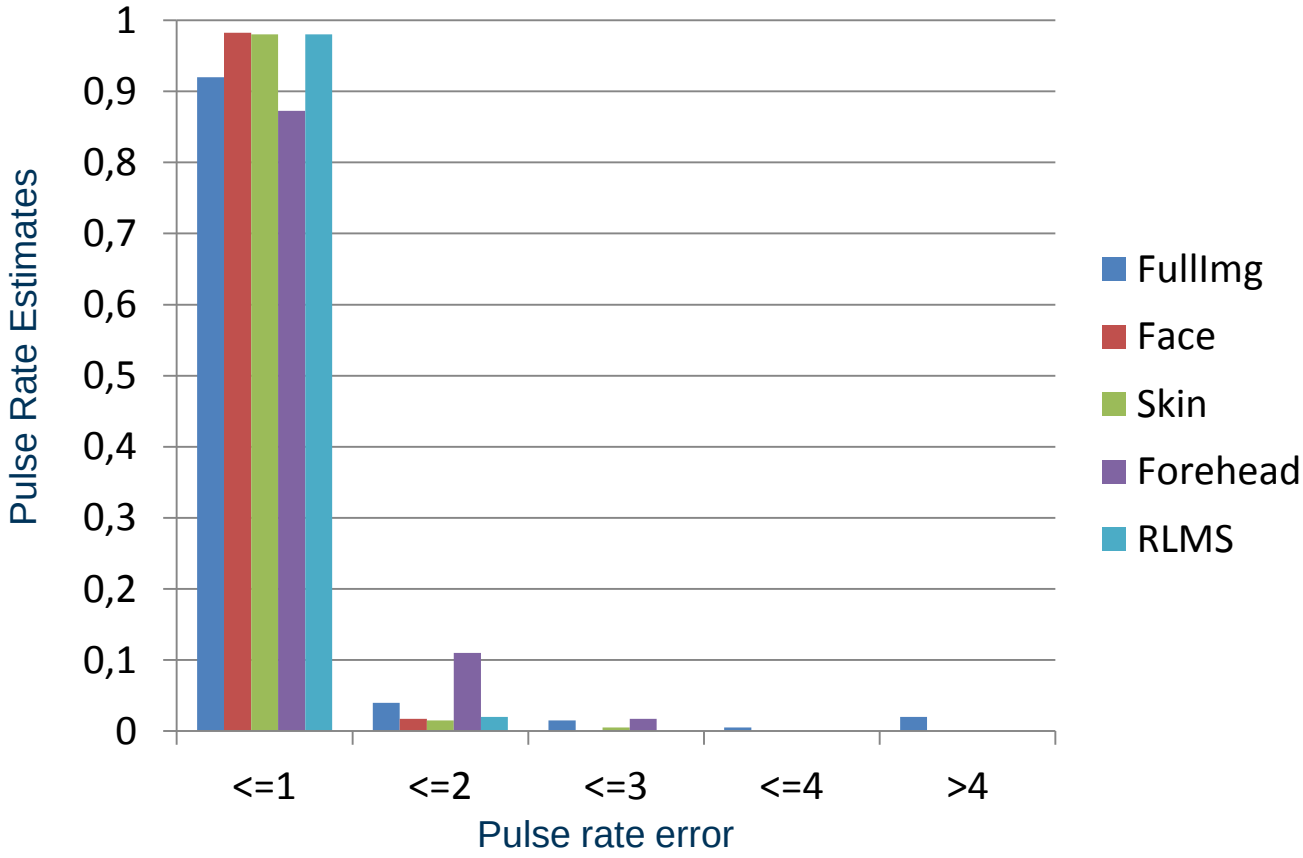
- One dimensional signal is required to perform FFT
- Using the green or the normalized green channel is common in literature
- Normalization: $Ch_{Norm} = \frac{Ch}{R + G + B}$



- Normalization reduces global illumination changes and increases influence of color hue

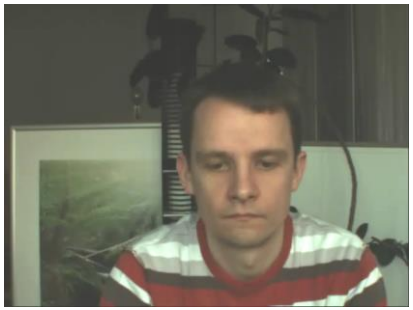
Pulse Measurement Error - Steady

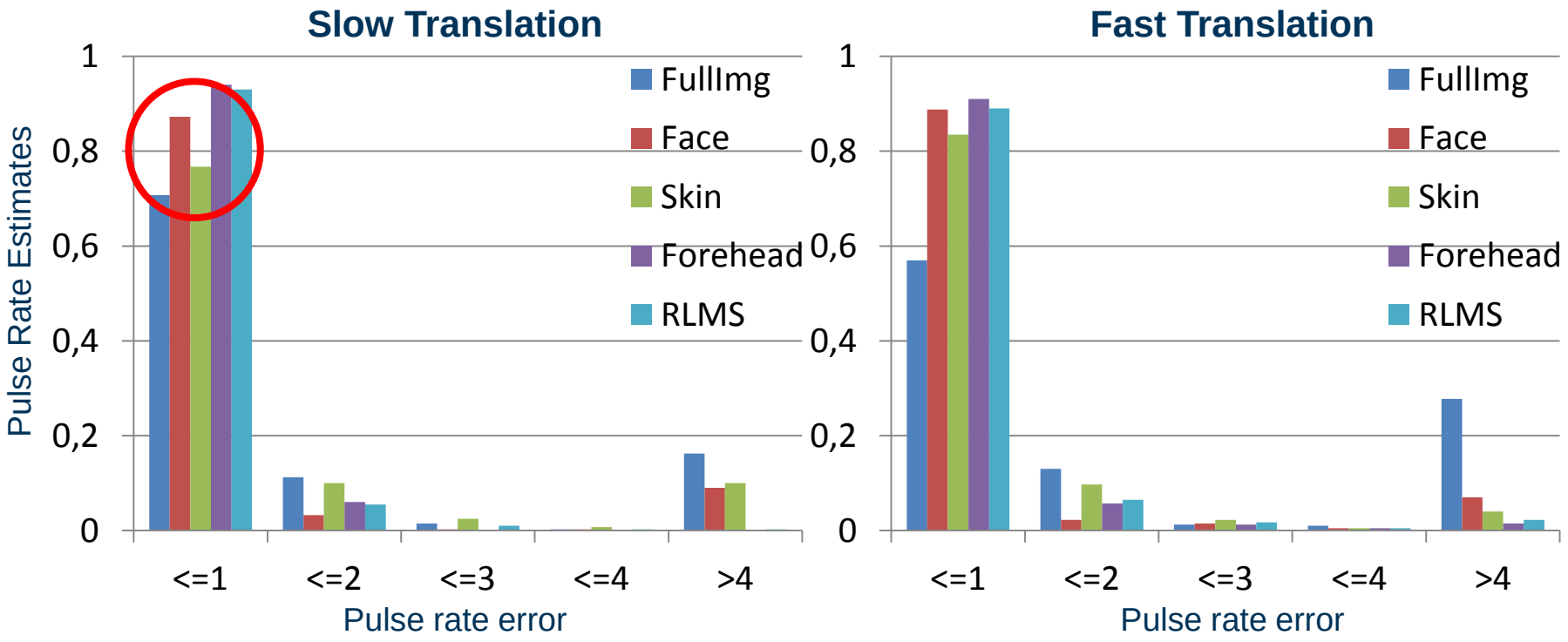
- Measuring pulse rate on the **Steady sequences** can be performed **very robust**
- **Almost perfect** measurements for all ROI extraction methods.



Pulse Measurement Error - Translation

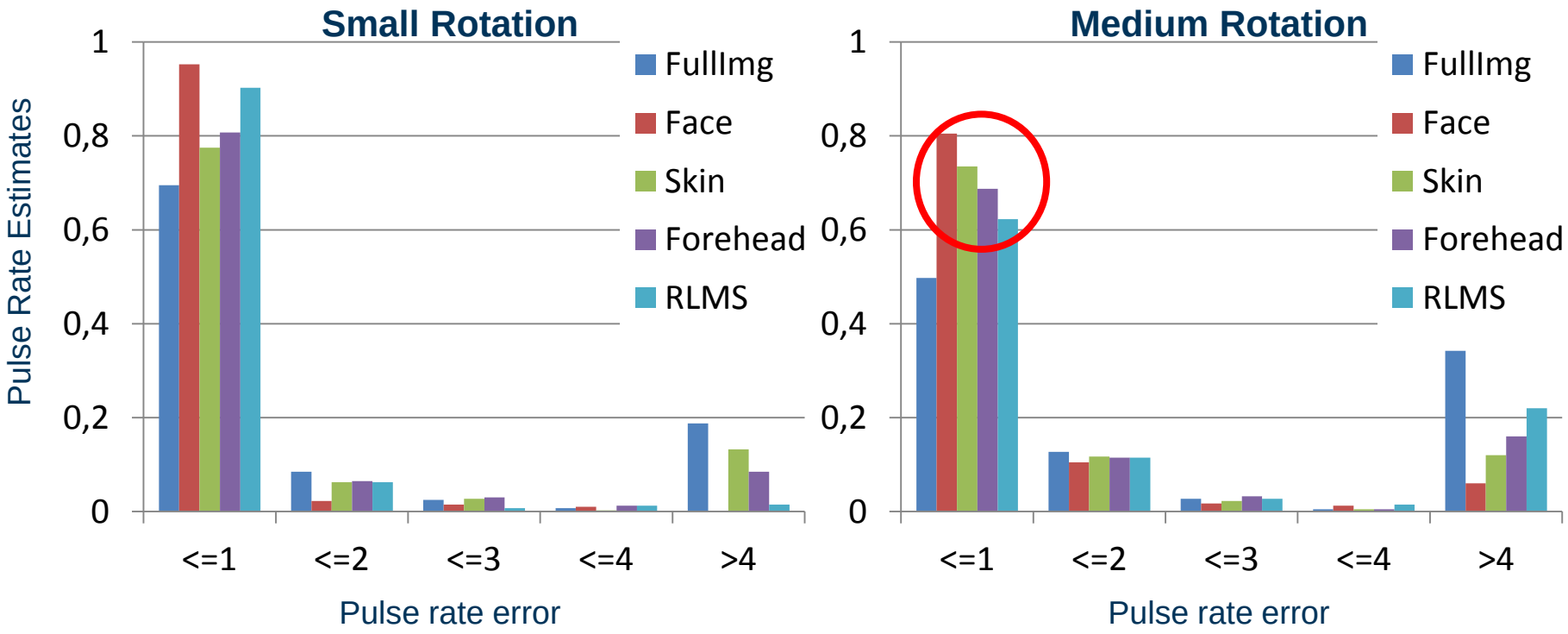
- **In plane** head translation handled **well** by all methods
- Slight advantage for RLMS and Forehead
- **Skin segmentation** very over time what introduces **noise**





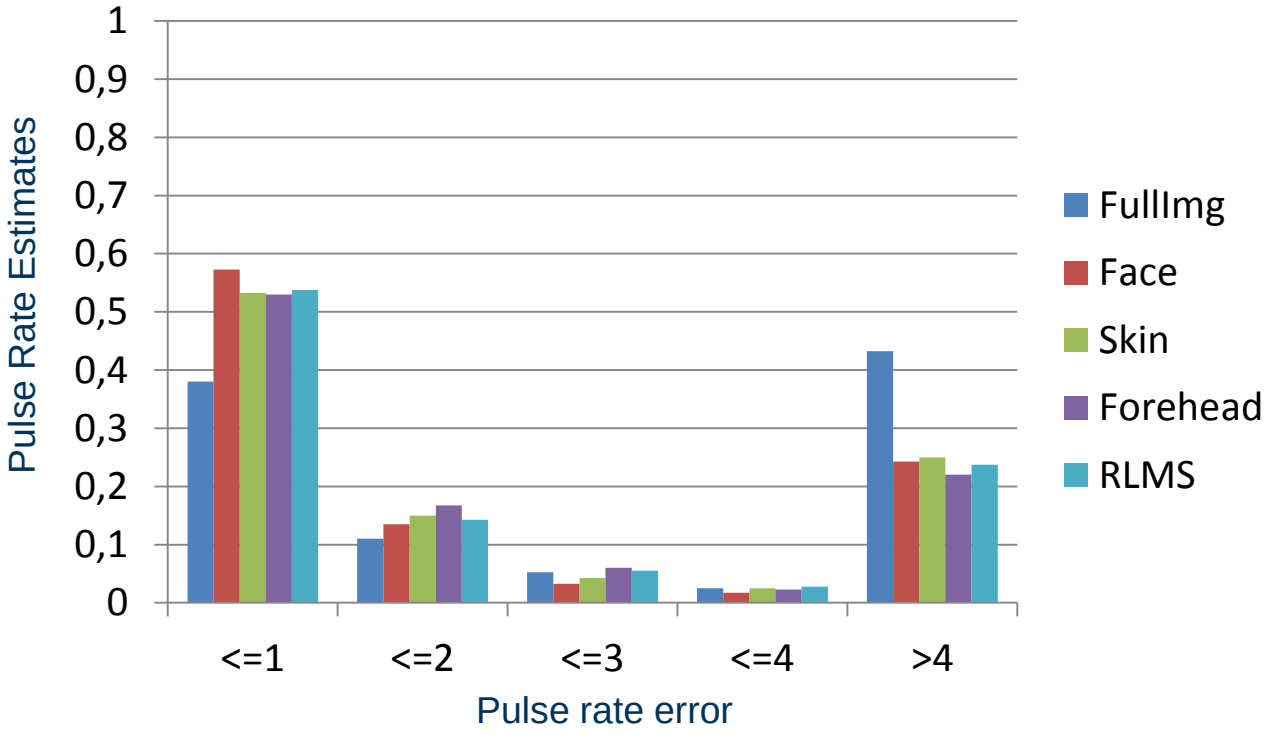
Pulse Measurement Error - Rotation

- **Surprisingly**, **face** tracking performs **better than RLMS**
- **RLMS** sometimes **fails** to reliably track face outline and introduces noise



Pulse Measurement Error - Talking

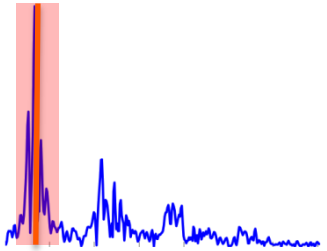
- **most challenging** in our dataset
- Pulse rate histograms are **bad** for all presented methods
- on a real robot **RLMS** can be favored, if coupled with a **talking classifier**

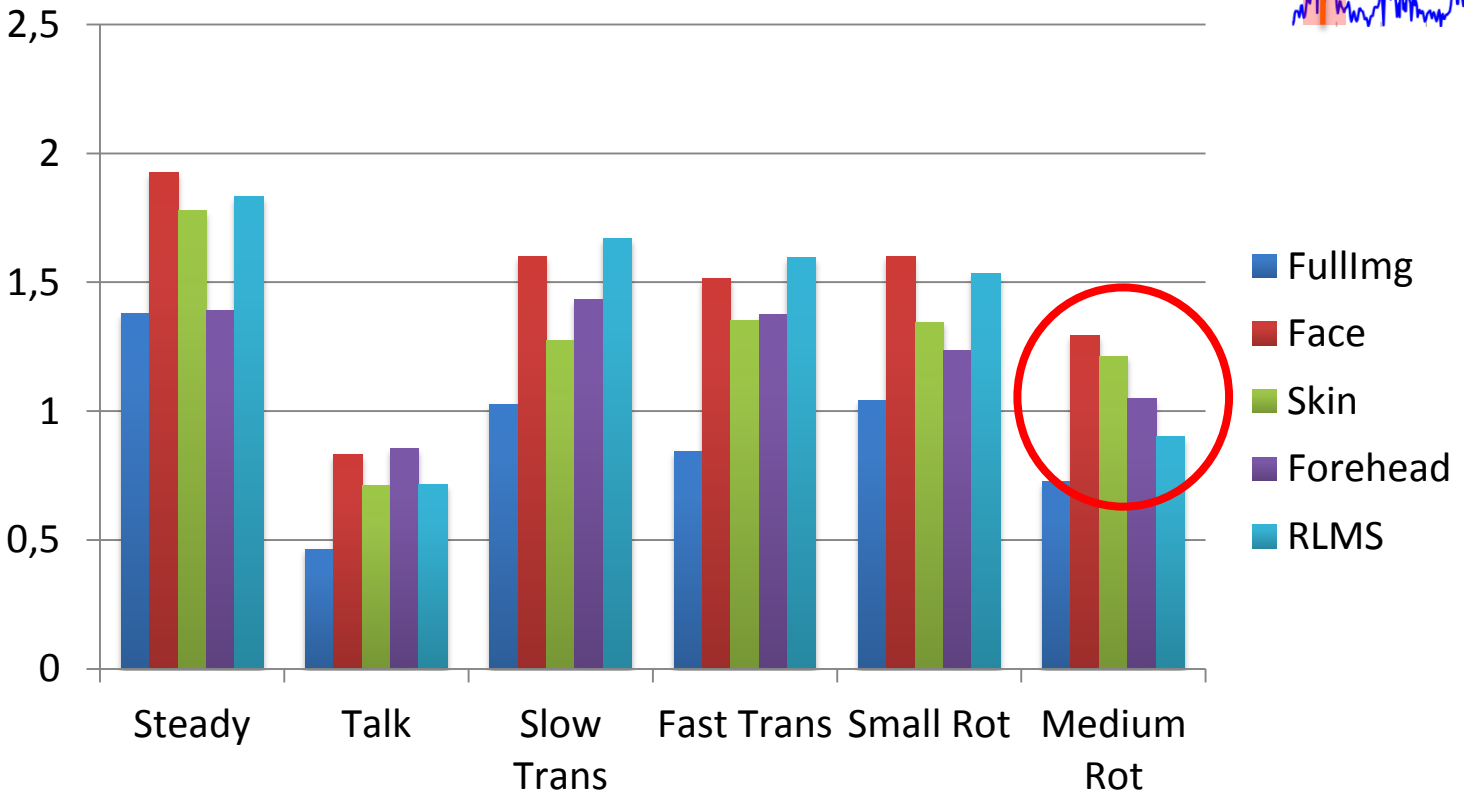


Signal to Noise Ratio

- RMSE is no good measure if error can become arbitrary large
- We evaluate sequences with Signal Noise Ratio and Deviation Histograms

$$SNR = \frac{\sum_{n=A}^{B} spect(n)}{\sum_{n=MinPulseBin}^{MaxPulseBin} spect(n) - \sum_{n=A}^{B} spect(n)}$$





Computational Complexity / Next Steps

- Robot equipped with an Intel Core i7-2640m 2.8 GHz (2 Cores)
- only one core used for runtime evaluation.

[ms]	FullImage	Face	Skin	Forehead	RLMS
ROI Extraction	0	10.4	10.4+57.4	12.9+0	12.9
Max FPS (hier nur ROI)	∞*	96.2	14.7	77.5	77.5
Post-processing with FFT	0.5				

Outlook / Summary

- Focus on the **region** used for **signal extraction** in more detail
- First attempts of extracting **single triangles from the RLMS mesh**
- Using the **RLMS model for initialization** of a face template, which is used for tracking afterward

Summary

- **Introduced new database** for evaluation of influence of motion for remote pulse measurement, which is **free to use**
- Presented **results** for different **SoA ROI extraction** methods
  - **Very similar** of **no motion** is present
  - **RLMS and face detection** are favored if **head motion** do occur
  - **RLMS** suited best for detecting **talking** sequences

Thank you !
Questions ?