Looking at People
CVPRW 2015

Spatio-temporal Analysis of RGB-D-T Facial Images for Multimodal Pain Level Recognition

Ramin Irani, Kamal Nasrollahi, C. Bahnsen, D.H. Lundtoft, T.B. Moeslund
Visual Analysis of People Lab, Aalborg University, Denmark

Marc O. Simon, Ciprian A. Corneanu, Sergio Escalera
Computer Vision Center, UAB, Barcelona, Spain

Tanja L. Pedersen, Maria-Louise Klitgaard, and Laura Petrini
Dept. of Communication and Psychology, Aalborg University, Denmark

Motivation

- We live longer =>
  pressure on the healthcare sector
- Suggested solution: Apply technology (somehow…)

- Automatic health assessment
  - Pain is a key indicator for many “medical conditions”
- Rehap@home
  - We are missing the therapist
  - Pain is a primary indicator

Research questions:
- Can we estimate pain via video analysis?
- Can multi-modal data help?

CVPR 2015

Methodology

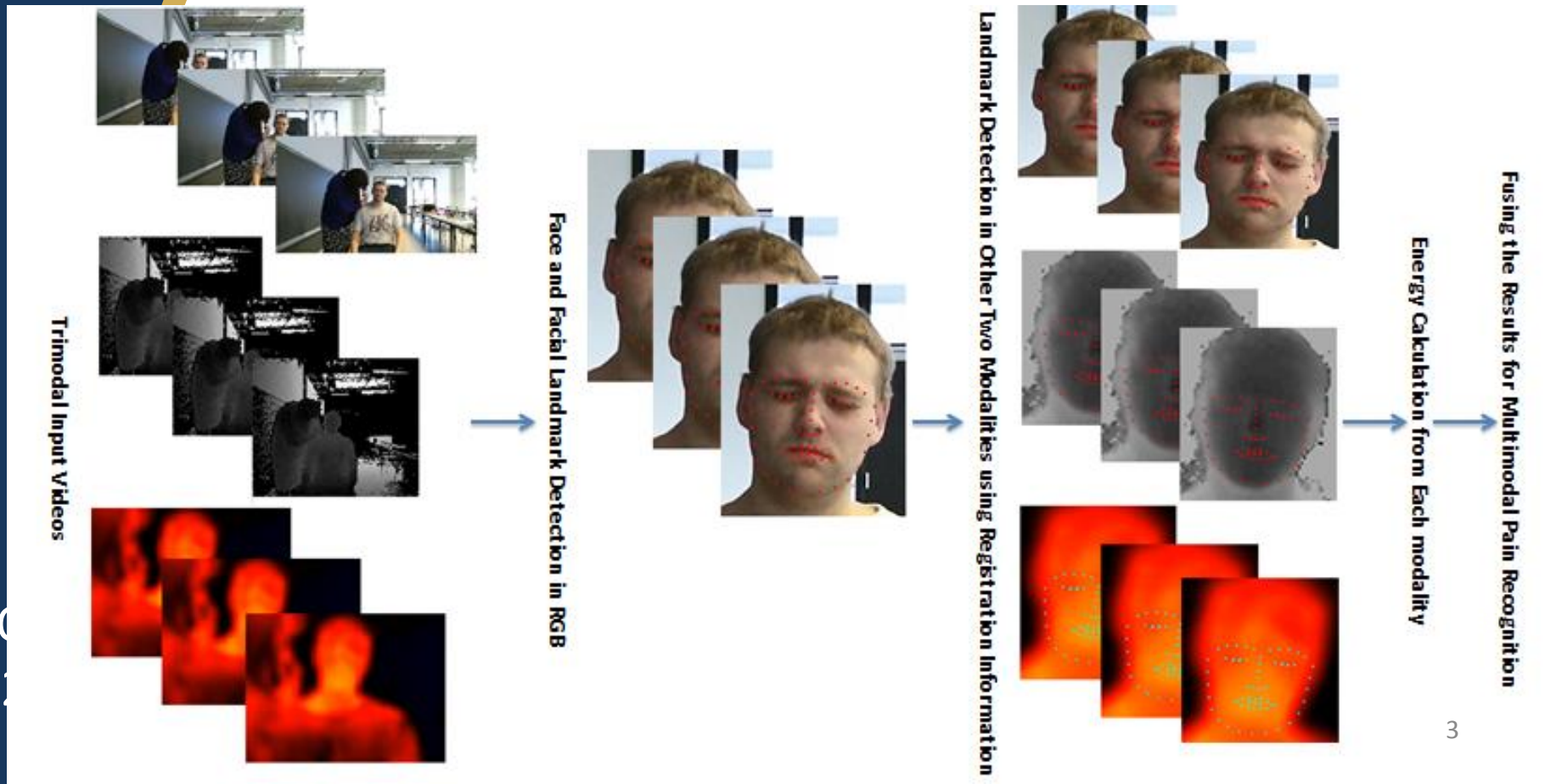
Face
-Detection
-Alignment
-Inpainting

→

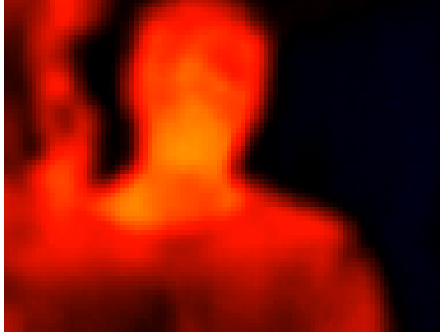
Spatiotemporal
Feature Extraction

→

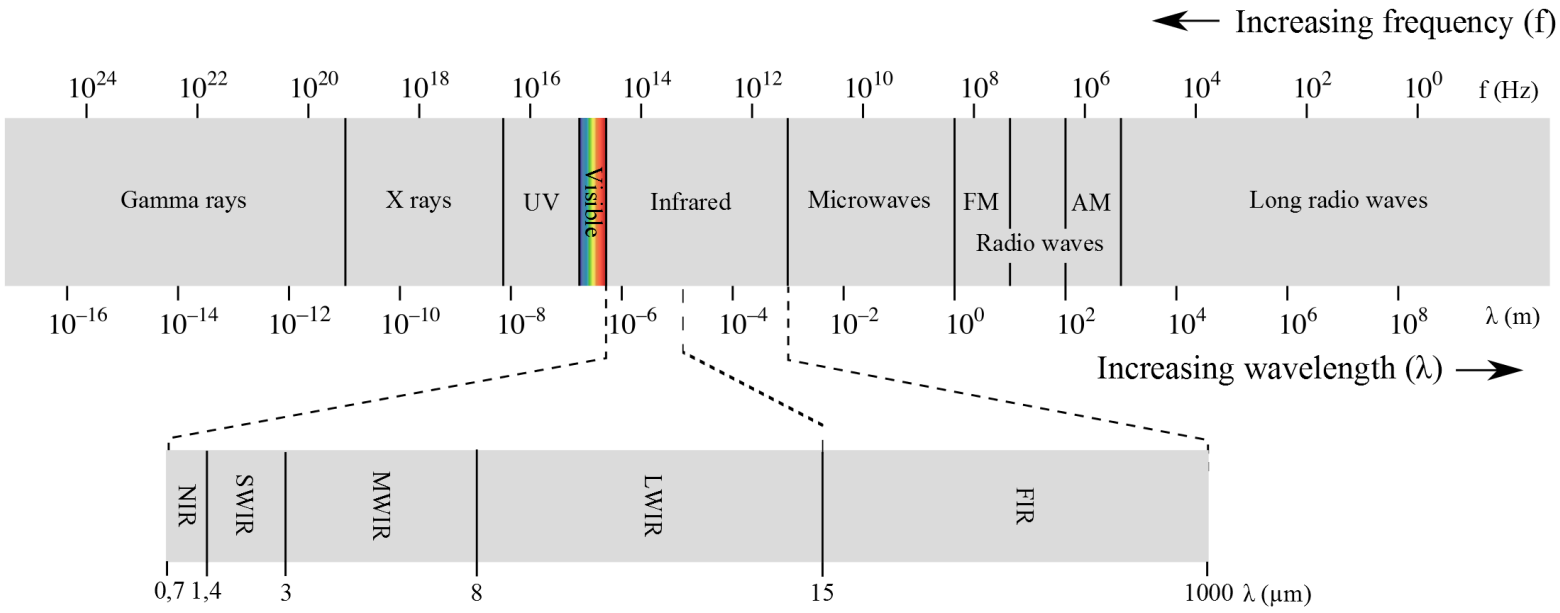
Pain
Recognition



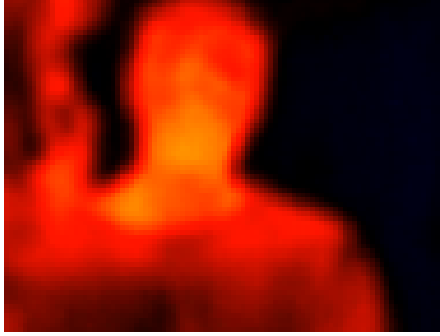
Thermal Imaging



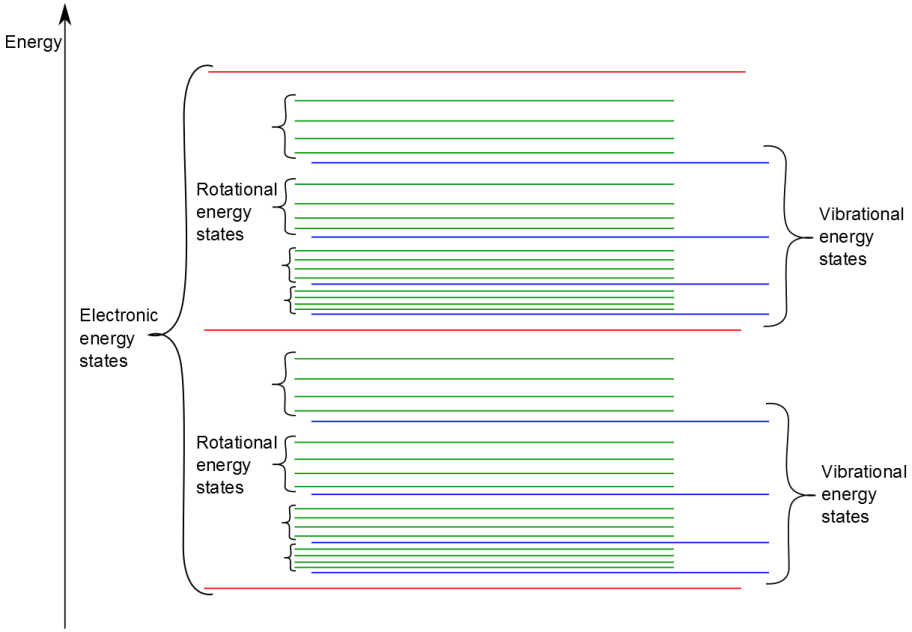
- Thermal cameras are sensitive to either mid-wavelength or long-wavelength infrared radiation

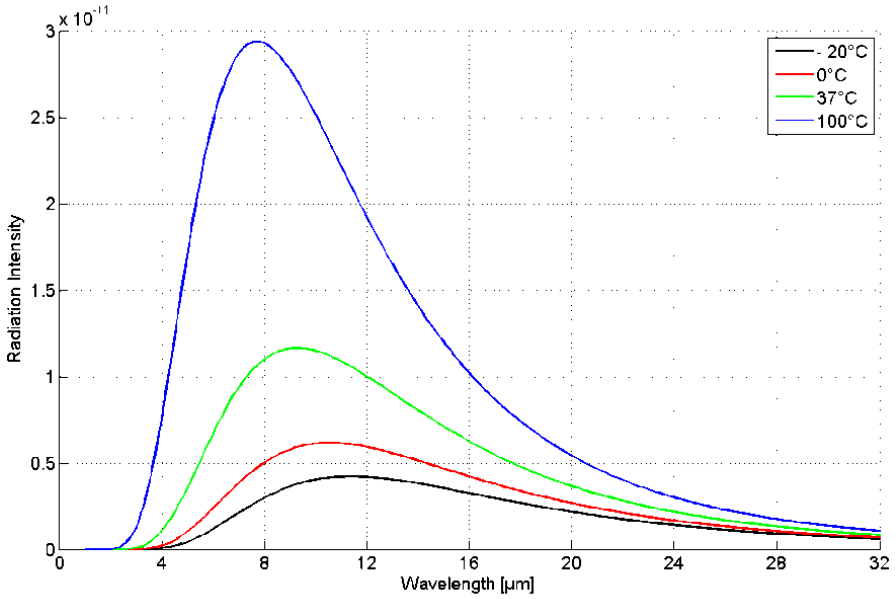


Thermal energy

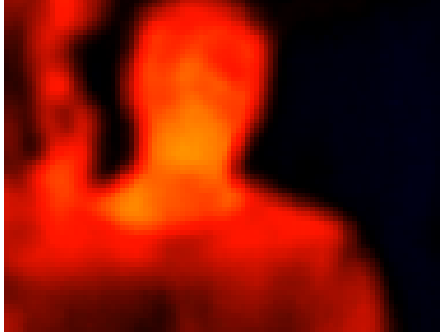


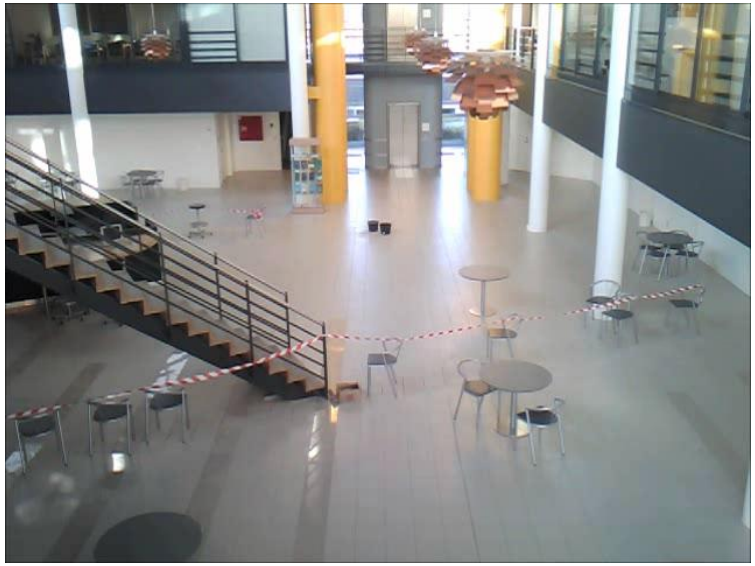
- Thermal radiation is caused by rotation and vibration in the molecules
- The intensity and dominating wavelength depend on the temperature (Planck's Law)





Thermal Imaging





Methodology

Landmark Detection in RGB

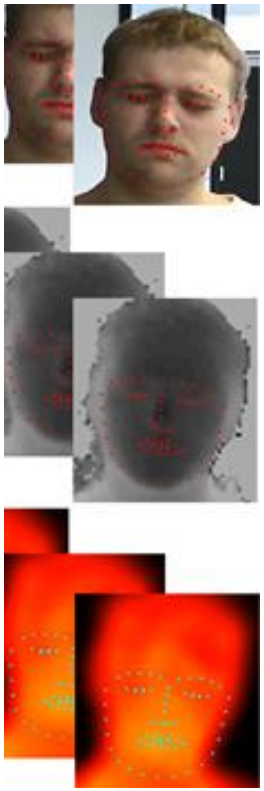
- The Viola & Jones in the first frame
- Landmarks are located inside the facial region by using the Supervised Descent Method (SDM)
  - An iterative shape model
  - Local texture for each landmark
- A custom implementation of SDM trained for the detection of 68 landmarks
- For training:
  - Training: LFPW, HELEN, AFW and IBUG datasets
  - Ground truth: 300 Faces In-The-Wild Challenge
- In the subsequent frames, the facial region is obtained from the previous frame geometry
- Applying SDM inside that region to estimate the new landmark locations.

CVPR 2015

X. Xiong and F. De la Torre. Supervised descent method and its applications to face alignment. CVPR'13

Methodology

Continue



Landmark detection in depth and thermal:

- Registration between the RGB and depth uses the built-in calibration tool of the Kinect for Windows 2.0 SDK.

- Registration of the thermal modality to RGB is performed via a custom-made multimodal checkerboard

CVPR 2015

Methodology

multimodal calibration

- Similar appearance in RGB and thermal modalities
- Two layers of cardboard
- Heat white layer and cool black layer before assembly
- Only few minutes calibration time before heating/cooling must be repeated

RGB

Thermal

CVPR 2015

Methodology

- Same as before for each modality:
  - Temporal alignment, warping, extract directional energy, histogram for each region, spatio-temporal features, weighting =>
  - Pain index: PI(t)
- Combine three modalities:

$$PI(t) = W_{\mathrm{RGB}}\, PI_{RGB}(t) + W_d\, PI_d(t) + W_t\, PI_t(t)$$

- Weights: 0.6, 0.35, and 0.05

CVPR 2015

Experiment and Results

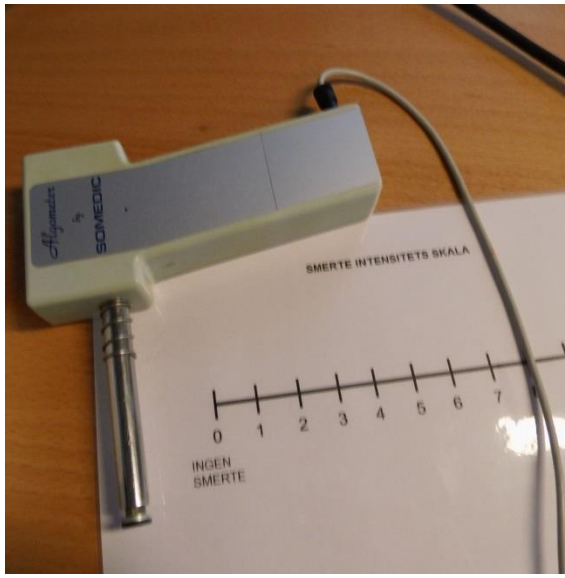
Setup and data

- New dataset
- Participants:
  - 12 healthy elderly volunteers (all females)
  - the ages of 66 and 90 years (mean age 73.6 years)
  - All subjects were pain-free and none of them had taken any analgesic or sedative for at least 48 hours prior to the experiment

CVPR 2015

Experiment and Results
Continue

- Hand-held pressure algometer was used to produce mechanical pressure
- Subjects' pain threshold learned (PDT)
- Subjects' pain self-reports were recorded using a numerical rating scale (NRS)
- The NRS ranges from 0 (no pain) to 10 (the worst pain you can imagine)



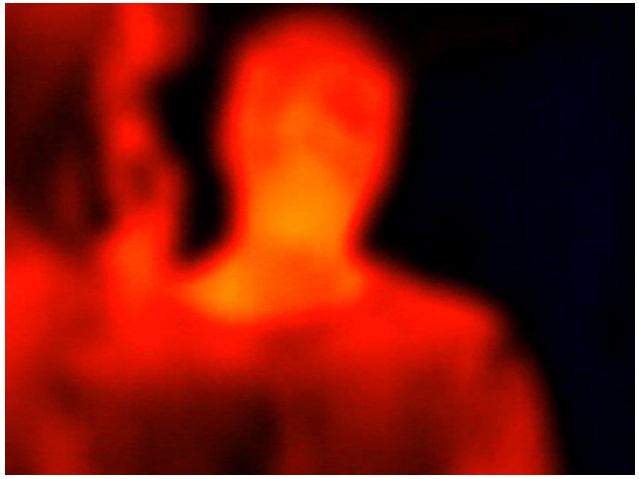
- Pain ground truth is calculated as:
  No-Pain: 0.2 X PDT,
  Light Pain: 1.10 X PDT,
  Moderate Pain: 1.30 X PDT,
  Strong Pain: 1.5 X PDT,

CVPR 2015

Experiment and Results

a

b



c

a) RBG (Kinect)
b) Depth (Kinect)
c) Thermal (AXIS Q1921)

(show video)

CVPR 2015

Experiment and Results discussion

- Comparing the results of the proposed system against the system of [*]:

Semantic Ground Truth	Pain Index Ground Truth	Number of Frames	System of [*] (in %)	Proposed System(in %)
No pain	1, 2	757	72	88
Weak	2,3,4,5	427	79	87
Strong	≥6	1204	76	76

[*] R. Irani, K. Nasrollahi, and T. B. Moeslund. Pain recognition using spatio-temporal oriented energy of facial muscles. In Computer Vision and Pattern Recognition Workshop, 2015

Conclusion

- The proposed system in this paper uses a spatio-temporal approach to extract facial energies
- Modalities: RGB, depth, and thermal
- Improving results over RGB by 6%
- The experimental results on a group of 12 elderly people showed that the proposed system can assist to detect the pain
- Future work:
  - More data
  - Better thermal data
  - Weighting/fusion

CVPR 2015

Thanks for your attention

Q & A

Work in progress =>

comments/suggestions are most welcome!

CVPR
2015