DEEP PERCEPTUAL MAPPING FOR THERMAL TO VISIBLE FACE RECOGNITION

M. Saquib Sarfraz and Rainer Stiefelhagen
Karlsruhe Institute of Technology
WHY THERMAL

• All objects emits thermal radiation

• Complete darkness, haze, rain etc.

• Even in daylight: Deep shadows, difficult lighting conditions
INFRARED SPECTRUM
INFRARED SPECTRUM

RESEARCH GAP
THE HEAT IS ON

- Cool is the new hot – **MWIR images**
  - Better spatial resolution, higher thermal contrast, distinguish smaller temperature differences, higher resolution images

- Cool but uncooled – **LWIR images**
  - Easy use, practical applications, much lower spatial resolution, noisy and hard to interpret images, larger research gap
THERMAL-VISIBLE FACE RECOGNITION

- Works in total darkness
- **Main Problem:** Wide modality and resolution difference
- Performance In a typical police mugshot matching

LWIR – LWIR = \(~ 90\%\)

Visible – LWIR = \(~ 35\%\)
WHAT HAS BEEN DONE

• Earlier work
  mostly NIR – visible

*Thermal*
• MWIR – visible: Klair & Jain 2013, similarity kernel subspace

• LWIR – Visible: state-of-the-art: Hu et al. 2015, Using PLS
DEEP PERCEPTUAL MAPPING (DPM)

Idea

- non-linear mapping between visible and thermal domain
- train a fully connected feed forward deep neural network
DEEP PERCEPTUAL MODEL

- learn a parametric non-linear perceptual mapping function

\[ H(x) = h^{(N)} = g(W^{(N)}h^{(N-1)} + b^{(N)}) \]

Objective function
- minimize perceptual difference

\[ \arg\min_{W,b} J = \frac{1}{M} \sum_{i=1}^{M} (x_i - t_i)^2 + \frac{\lambda}{N} \sum_{k=1}^{N} (\|W^{(k)}\|_F^2 + \|b^{(k)}\|_2^2) \]
• Dense SIFT vectors @ two scales – 408 vectors/image
• PCA + position embedding
• ~1 million vectors to train the model
• captures the differing local regions perceptual differences
• Loss minimized by standard back projection using SGD
FACE RECOGNITION

- Concatenate vectors
- **Gallery**: mapped visible data \((\overline{x}_i)\)
- **Probes**: matched directly \((t_j)\)

- **Recognition**: one matrix-vector multiplication

\[
d(\overline{x}_i \cdot t_j) = \overline{x}_i \cdot t_j \quad \forall i = 1, 2, \ldots, G
\]
EXPERIMENTS

- UND-X1 Data (LWIR /Visible images)
- Baseline: Same features without the mapping

We analyze

- Effect of modality gap
- Identification evaluation
- Verification evaluation
**UN D-X1 DATA**

**Problems:** Wide Modality difference, resolution difference, stand-off distance, facial appearance (Pose, expressions, ambient lighting etc.)

- Resolution: Visible (1600x1200), LWIR (312 x 239)
- 4584 images, 82 subjects (equal split)
PREPROCESSING IS IMPORTANT

- Zero-mean normalization
- Median filtering (for dead-pixel removal)
- Difference of Gaussian filtering
## Results

### Effect of Modality gap: Performance with 1 Gallery image/subject

<table>
<thead>
<tr>
<th></th>
<th>Thermal - Thermal</th>
<th>Thermal-Visible</th>
<th>Thermal-Vis (via DPM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>89.47 %</td>
<td>30.36 %</td>
<td>55.36 %</td>
</tr>
</tbody>
</table>

~60% performance drop is bridged by 42%
## IDENTIFICATION COMPARISON

<table>
<thead>
<tr>
<th></th>
<th>Gallery size: # of visible images/subject</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1/subject</td>
</tr>
<tr>
<td>Baseline (same features w/o DPM)</td>
<td>30.3</td>
</tr>
<tr>
<td>Choi et al. 2012</td>
<td>-</td>
</tr>
<tr>
<td>Hu et al. 2015</td>
<td>46.3</td>
</tr>
<tr>
<td>DPM – Shallow (ours)</td>
<td>50.6</td>
</tr>
<tr>
<td>DPM (ours)</td>
<td><strong>55.3</strong></td>
</tr>
</tbody>
</table>
IDENTIFICATION CMC

- Baseline - 1 gallery/sub: 30.4%
- Baseline - all gallery/sub: 52.3%
- DPM - 1 gallery/sub: 55.4%
- DPM - all gallery/sub: 83.7%
VERIFICATION

ROC: 1792 genuine and 71680 imposter attempts
(2 visible images/sub in gallery)
COMPUTATIONAL TIME

- Training: CPU multicore, 1-1.5 hours on MATLAB.
- Test time: identifying one probe only takes 35 ms
- Fast and capable of running in real-time
CONCLUSIONS

• Thermal-visible face matching desired but difficult, large research gap

• DPM is effective and has benefits for related problems e.g., domain adaptation, remote sensing and other general CV problems

• Computationally efficient

• Face recognition can work in complete darkness
Deep Perceptual Mapping for Thermal-Visible Face Recognition

Contact: M. Saquib Sarfraz, cv:hci Lab, KIT, Germany
Web: https://cvhci.anthropomatik.kit.edu/~ssarfraz