Recent Advances and Challenges in Facial Micro-Expression Analysis

Spotting ME Sequences

John See  Multimedia University, Malaysia
Moi-Hoon Yap  Manchester Metropolitan University, UK
Su-Jing Wang  Chinese Academy of Sciences, China
Jingting Li  Chinese Academy of Sciences, China
Sze-Teng Liong  Feng Chia University, Taiwan
Outline

- Related Works
- Main Directional Maximal Difference Analysis
- Spatio-temporal Fusion for Macro- and Micro-expression Spotting in Long Video Sequences
- Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting
- Result Evaluation Method per Interval
Outline

- **Related Works**
- Main Directional Maximal Difference Analysis
- Spatio-temporal Fusion for Macro- and Micro-expression Spotting in Long Video Sequences
- Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting
- Result Evaluation Method per Interval
Related works

**Feature Extraction**
- LBP-χ²
- HOG-χ²
- HOOF
- MDMD
- Etc.

**Early Fusion**
- Global feature for entire face

**Feature Difference in temporal interval based on ME duration**

**Classification**
- SVM
- Adaboost model
- CNN

Small data size
Outline

- Related Works
- **Main Directional Maximal Difference Analysis**
  - Spatio-temporal fusion for Macro- and Micro-expression Spotting in Long Video Sequences
  - Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting
- Result Evaluation Method per Interval
Main Directional Maximal Difference (MDMD) Analysis

**Contribution:**

- MDMD uses the magnitude of maximal difference in the main direction of optical flow as a feature.
- Based on block-structured facial regions, MDMD obtains more accurate features of the movement of expressions.
- This method obtains both the temporal and spatial locations of facial movements.
- **Baseline method** for Micro-expression Grand Challenge (MEGC) 2020 - spotting micro-expressions and macro-expressions on long videos.


Main Directional Maximal Difference (MDMD) Analysis

- Pre-process

Main Directional Maximal Difference (MDMD) Analysis

- Feature Extraction

\[
\{(\rho_M^{HC}, \theta_M^{HC})\} = \{(\rho_M^{HC}, \theta_M^{HC}) | \theta_M^{HC} \in \Theta\}
\]

\[
\{(\rho_M^{HT}, \theta_M^{HT})\} = \{(\rho_M^{HT}, \theta_M^{HT}) | \theta_M^{HT} \} \text{ and } (\rho_M^{HC}, \theta_M^{HC})
\]

are two different vectors of the same point in \( F_{i-k} \)

\[
d = \frac{3}{g} \sum \max \left\{ \rho_M^{HC} - \rho_M^{HT} \right\}
\]


Main Directional Maximal Difference (MDMD) Analysis

- Block-structured analysis

\[ \bar{d}^i = \frac{1}{s} \sum_{s} \max \{d^i_j\} \]

\[ r^i = \bar{d}^i - \frac{1}{2} (\bar{d}^{i-k+1} + \bar{d}^{i+k-1}) \]

\[ \text{threshold} = r_{mean} + p \times (r_{max} - r_{mean}) \]

where

\[ \text{mean} = \frac{1}{n-2k} \sum_{n-k}^{i=k+1} r^i \]

and

\[ r_{max} = \max_{n-k} r^i. \]


Outline

- Related Works
- Main Directional Maximal Difference Analysis
- Spatio-temporal fusion for Macro- and Micro-expression Spotting in Long Video Sequences
- Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting
- Result Evaluation Method per Interval

Jingting LI, Institute of Psychology, Chinese Academy of Sciences
Spatio-Temporal fusion for Macro- and Micro-expression Spotting in Long Video Sequences

Contribution:

- We address the head motion problem in a simple way. We select the nose region as a standard global vector that contains only head motion. The local optical flow field is obtained by performing the operator of difference between superposition of optical field and standard global movement vector.

- Second, we propose a Spatio-temporal feature fusion matrix which describes spatial and temporal information by row and column relationship. A specific pattern related to magnitude and angle is extracted from the matrix. We denoted it as SP-pattern, which contains all the information from a micro-expression interval. We can obtain onset, apex, and offset according to the SP-pattern.

- Third, we use a multi-scale filter to remove high frequency noise and preserve crests of different intensities. In order to achieve good performance on both macro-expression and micro-expression, we comprehensively analyze information at different scales.

- Won the first place of MEGC2020

Spatio-Temporal fusion for Macro- and Micro-expression Spotting in Long Video Sequences

Preprocessing  Feature extraction  Spotting process

<table>
<thead>
<tr>
<th>Facial region</th>
<th>ROI</th>
<th>AU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eyebrows</td>
<td>1, 2, 3, 4, 5, 6</td>
<td>1, 2, 4, 5, 6, 7, 9</td>
</tr>
<tr>
<td>Nose</td>
<td>7, 12</td>
<td>11</td>
</tr>
<tr>
<td>Mouth</td>
<td>8, 9, 10, 11</td>
<td>10 12 14 15 17 20 23 24</td>
</tr>
</tbody>
</table>

Spatio-Temporal fusion for Macro- and Micro-expression Spotting in Long Video Sequences

Preprocessing

Feature extraction

Spotting process

Outline

- Related Works
- Main Directional Maximal Difference Analysis
- Spatio-temporal Fusion for Macro- and Micro-expression Spotting in Long Video Sequences
- **Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting**
- Result Evaluation Method per Interval
Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

**Contribution:**
- Local Temporal Pattern (LTP)
- A late spatial-and-temporal fusion
- Data augmentation by Hammerstein Model
- A novel result evaluation method and metric


Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

- Region of Interest selection

<table>
<thead>
<tr>
<th>Facial region</th>
<th>Related AU</th>
<th>12 ROI index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eyebrows</td>
<td>1, 2, 4</td>
<td>1, 4, 5, 6, 7, 10</td>
</tr>
<tr>
<td>Nose</td>
<td>NaN</td>
<td>11, 14</td>
</tr>
<tr>
<td>Mouth</td>
<td>10, 12, 14, 15, 17, 25</td>
<td>32, 35, 38, 41</td>
</tr>
</tbody>
</table>

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Local temporal pattern calculation:

Why S-pattern can be used for classification?

Same S-pattern for ME for different emotions and different ROIs

Different variations for ME (S-pattern) and non-ME region (Other LTPs)

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Feature annotation (S-pattern and non S-pattern) + SVM (LOSubOCV) = ROIs with S-pattern are recognized

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

**Late spatial and temporal fusion**

- **Local classification**
- **Local qualification per ROI**
- **Spatial Fusion from ROIs to entire face**
- **Merge process**
- **Final Spotting result**

**Graphs:**
- **ROI 4 (eyebrows)**
- **ROI 11 (nose)**
- **ROI 32 (mouth)**

---

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Late spatial and temporal fusion

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Late spatial and temporal fusion

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Late spatial and temporal fusion

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

**Late spatial and temporal fusion**

- **Local classification**
- **Local qualification per ROI**
- **Spatial Fusion from ROIs to entire face**
- **Merge process**
- **Final Spotting result**

---

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Late spatial and temporal fusion

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

- **Hammerstein Model**
  - A discrete-Time Model of Electrically Stimulated Muscle.
  - Response to the physique explanation of pattern synthetize.

\[
\frac{Y(s)}{V(s)} = \frac{s}{1 + \alpha s + \beta s^2}
\]

Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Data Augmentation by Hammerstein Model

Hammerstein Model: 
\[(p, \alpha, \beta, E_H)\]

System identification

Generation

S-patterns, which are wrongly labeled, can be filtered by $E_H$

The $(\alpha, \beta)$ distribution allows to define the $(\alpha, \beta)$ value range for synthesizing reliable S-patterns by Hammerstein model.
Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting

Data augmentation using Hammerstein model
→ Better classification performance

Outline

- Related Works
- Main Directional Maximal Difference Analysis
- Spatio-temporal Fusion for Macro- and Micro-expression Spotting in Long Video Sequences
- Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting
- Result Evaluation Method per Interval
Result Evaluation Method per Interval

- MEGC 2019, 2020 spotting task

What we want:
- Avoid the inaccuracy of annotation
- Less the true negatives

Per interval
True positive in one video:
\[
\frac{W_{\text{spotted} \cap W_{\text{groundTruth}}}}{W_{\text{spotted} \cup W_{\text{groundTruth}}}} \geq k
\]
- where \( k = 0.5 \)

Result Evaluation Method per Interval

- Suppose that in one video
  - # of Ground truth (ME) interval: \( m \)
  - # of Spotted interval: \( n \)
  - # of TP interval: \( a \)

\[
\begin{align*}
W_{\text{groundTruth}} &\quad W_{\text{groundTruth}} \\
W_{\text{spotted}} &\quad W_{\text{spotted}} \\
W_{\text{spotted}} &\quad W_{\text{spotted}}
\end{align*}
\]

- Evaluation for entire database
  - \( \text{recall, TPR} = \frac{\sum_{i=1}^{V} a_i}{\sum_{i=1}^{V} m_i} = \frac{A}{M} \), \( \text{precision} = \frac{\sum_{i=1}^{V} a_i}{\sum_{i=1}^{V} n_i} = \frac{A}{N} \).
  - Imbalanced sample distribution: \( \text{F1-score} = \frac{2 \times (\text{recall} \times \text{precision})}{\text{recall} + \text{precision}} \).
  - Where \( a_i \) is the true positive, \( m_i \) and \( n_i \) are ME amount and spotted interval in \( i_{th} \) video respectively.


Jingting Li, FAST, IETR/CENTRALESUPELEC
Conclusion

- Main Directional Maximal Difference Analysis
- Spatio-temporal Fusion for Macro- and Micro-expression Spotting in Long Video Sequences
- Local Temporal Pattern and Data Augmentation for Micro-Expression Spotting
- Result Evaluation Method per Interval
## Perspectives

### Data augmentation for micro-expression spotting
- More micro-expression databases
- Synthesizing ME samples or features

### Consistency of metric and comparison
- F1-score and per interval
- Participating into challenges

### Micro-expression spotting applications
- Spotting in-the-wild
- Spotting in real time
- Combination with real application
- Fusion of macro- and micro-expression spotting
Thanks for your attention😊

Any questions?
lijt@psych.ac.cn