Multi-Instance Hidden Markov Model For Facial Expression Recognition
Chongliang Wu, Shangfei Wang, Qiang Ji
• Two approaches for facial expression recognition:
  ➢ Frame based
    • Recognize expressions from frames
    • Require pre-identified apex frames and its labels for training
  ➢ Sequence based
    • Recognize expressions from image sequences
    • Require pre-segmented sequences (neutral to apex) and their labels

Both approaches require preprocessing before performing facial expression recognition
Propose a sequence based facial expression recognition that

- Does not need pre-segmentation
- Only needs sequence level label. No segment label is needed
- Identify the label for each sequence and localize the expression segments

**Multiple instance learning (MIL) + Hidden Markov model (HMM)**
Divide training samples into bags

- Positive bag: at least one positive instance
- Negative bag: negative instances only

Training: given the bags \( X \) and their labels \( Y \), train an instance classifier \( Y_n = f(X_n, \theta) \) by maximizing the conditional log likelihood of the bag labels

\[
\log P(Y|X) = \sum_{X \in B^+} \log P(Y = 1|X, \theta) + \sum_{X \in B^-} \log P(Y = 0|X, \theta)
\]

Testing: For query bag \( X = \{X_n\} \), estimate the label \( Y_n \) for each instance \( X_n \) using the instance classifier and then the label of the bag as

\[
Y = \arg \max_n p(Y_n|X_n)
\]
MIL-HMM for Expression Recognition

- Construct bags and bag instances
  - Each sequence $S_i$ is a bag
  - For each sequence, divide it into segments $S_{ij}$ to form instances

- Within the MIL framework, train a HMM instance classifier to recognize expression for each segment

- During testing, given a query sequence $S$, divide it into segments, classify each segment using HMM, and determine expression label for $S$
1. Construct positive and negative sequences bags

2. Use N-cut algorithm divide each sequence into segments to produce instances for each bag
Given training sequences $S_i$ and its label $Y_i$, where $S_i=\{S_{ij}\}$ and $S_{ij}$ is the jth segment for ith sequence

Train the HMM instance classifier within the MIL framework by maximizing the CLL of the bag labels to solve for HMM parameters via gradient descent using BFGS method

For multi-class classification, one MIL-HMM model is constructed for each class.
Given query sequence $S_i$

1. Determine the label for each segment

$$Y_{ij} = \arg \max_{Y_{ij} = \{1, 2, \ldots, k\}} p(Y_{ij} | S_{ij})$$

2. Determine the label for the sequence

$$Y_i = \arg \max_j p(Y_{ij} | S_{ij})$$
Experiments

- **Extended Cohn-Kanade (CK+) Database**
  - 7 expression categories
  - 327 sequences
  - 118 subjects

- **UNBC-McMaster Shoulder Pain Expression Archive Database**
  - Pain expression
  - 149 sequences out of 200 selected
  - 25 subjects

- **Leave one subject out cross validation**
## Results

### CK+ Database

<table>
<thead>
<tr>
<th>Expression</th>
<th>Accuracy</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>angry</td>
<td>100%</td>
<td>1</td>
</tr>
<tr>
<td>contempt</td>
<td>98.15%</td>
<td>0.9714</td>
</tr>
<tr>
<td>disgust</td>
<td>97.18%</td>
<td>0.9558</td>
</tr>
<tr>
<td>fear</td>
<td>97.33%</td>
<td>0.9583</td>
</tr>
<tr>
<td>happiness</td>
<td>99.52%</td>
<td>0.9927</td>
</tr>
<tr>
<td>sadness</td>
<td>100%</td>
<td>1</td>
</tr>
<tr>
<td>surprise</td>
<td>97.59%</td>
<td>0.9652</td>
</tr>
<tr>
<td>average</td>
<td>98.54%</td>
<td>0.9776</td>
</tr>
</tbody>
</table>

### UNBC-McMaster Pain Database

<table>
<thead>
<tr>
<th>Classified as</th>
<th>positive</th>
<th>negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confusion Matrix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>positive</td>
<td>39</td>
<td>18</td>
</tr>
<tr>
<td>negative</td>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>85.23%</td>
<td></td>
</tr>
<tr>
<td>F1-score</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.78</td>
<td></td>
</tr>
</tbody>
</table>
Comparison with others on pain database

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accuracy</th>
<th># subjects</th>
<th># sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIL-HMM</td>
<td>85.23%</td>
<td>25</td>
<td>149</td>
</tr>
<tr>
<td>MS-MIL (K. Sikka et al. [1])</td>
<td>83.7%</td>
<td>23</td>
<td>147</td>
</tr>
<tr>
<td>P. Lucey et al. [9]</td>
<td>80.99%</td>
<td>20</td>
<td>142</td>
</tr>
<tr>
<td>A.B. Ashraf et al. [10]</td>
<td>68.31%</td>
<td>20</td>
<td>142</td>
</tr>
<tr>
<td>A.B. Ashraf et al. [10]</td>
<td>81.21%</td>
<td>21</td>
<td>84</td>
</tr>
</tbody>
</table>
Sample on CK+ Database

Segment Localization
Conclusion

• Method for sequence-based expression recognition
  ➢ MIL + HMM
  ➢ No prior sequence segmentation
  ➢ Only sequence level label is required
  ➢ Identify expression label for each sequence as well as localize the expression segments

• Experiments on CK & UNBC-McMaster Database
  ➢ Achieve excellent performance in recognition accuracy
  ➢ Outperform state of the art methods
THANKS!