中国科学技术大学
University of Science and Technology of China

Rensselaer

Multi-Instance Hidden Markov Model For Facial Expression Recognition

Chongliang Wu, Shangfei Wang, Qiang Ji

Motivation

- 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

Motivation

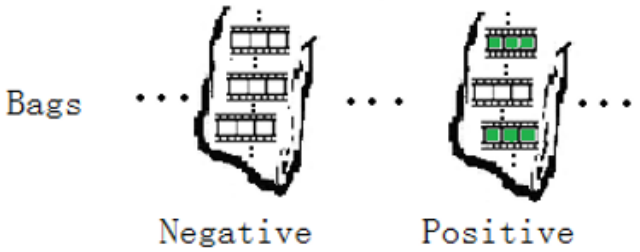
Propose a sequence based facial expression recognition that

- 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)

Multi-Instance Learning (MIL)

- 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 instances X_n using the instance classifier and then the label of the bag as

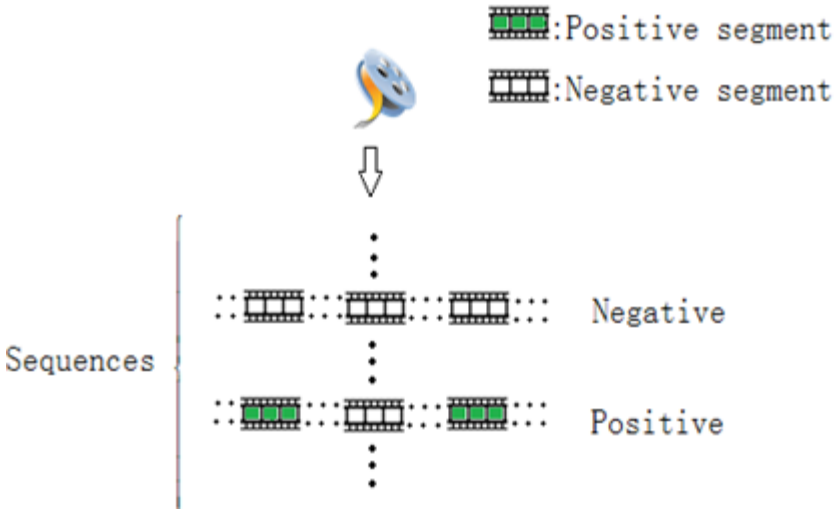
$$Y = arg \max_n \; p(Y_n|X_n)$$

MIL-HMM for Expression Recognition

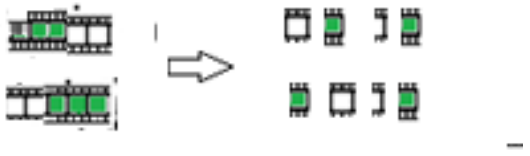
- Construct bags and bag instances
  - Each sequence S_i is a bag
  - For each sequence, divide it into segments S_{ij} to form instances
- Within the MIL framework, train a HMM instance classifier to recognize expression for each segment
- During testing, given a query sequence S, divide it into segments, classify each segment using HMM, and determine expression label for S

Construct bags and instances

1. **Construct positive and negative sequences bags**



2 **Use N-cut algorithm divide each sequence into segments to produce instances for each bag**



Joint MIL-HMM Training

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.

Testing

Given query sequence S_i

1. Determine the label for each segment

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

2. Determine the label for the sequence

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

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

Expression	Accuracy	F1-score
angry	100%	1
contempt	98.15%	0.9714
disgust	97.18%	0.9558
fear	97.33%	0.9583
happiness	99.52%	0.9927
sadness	100%	1
surprise	97.59%	0.9652
average	98.54%	0.9776

UNBC-McMaster Pain Database

Classified as			
Confusion Matrix		positive	negative
	positive	39	18
	negative	4	88
Accuracy	85.23%		
F1-score	0.78		

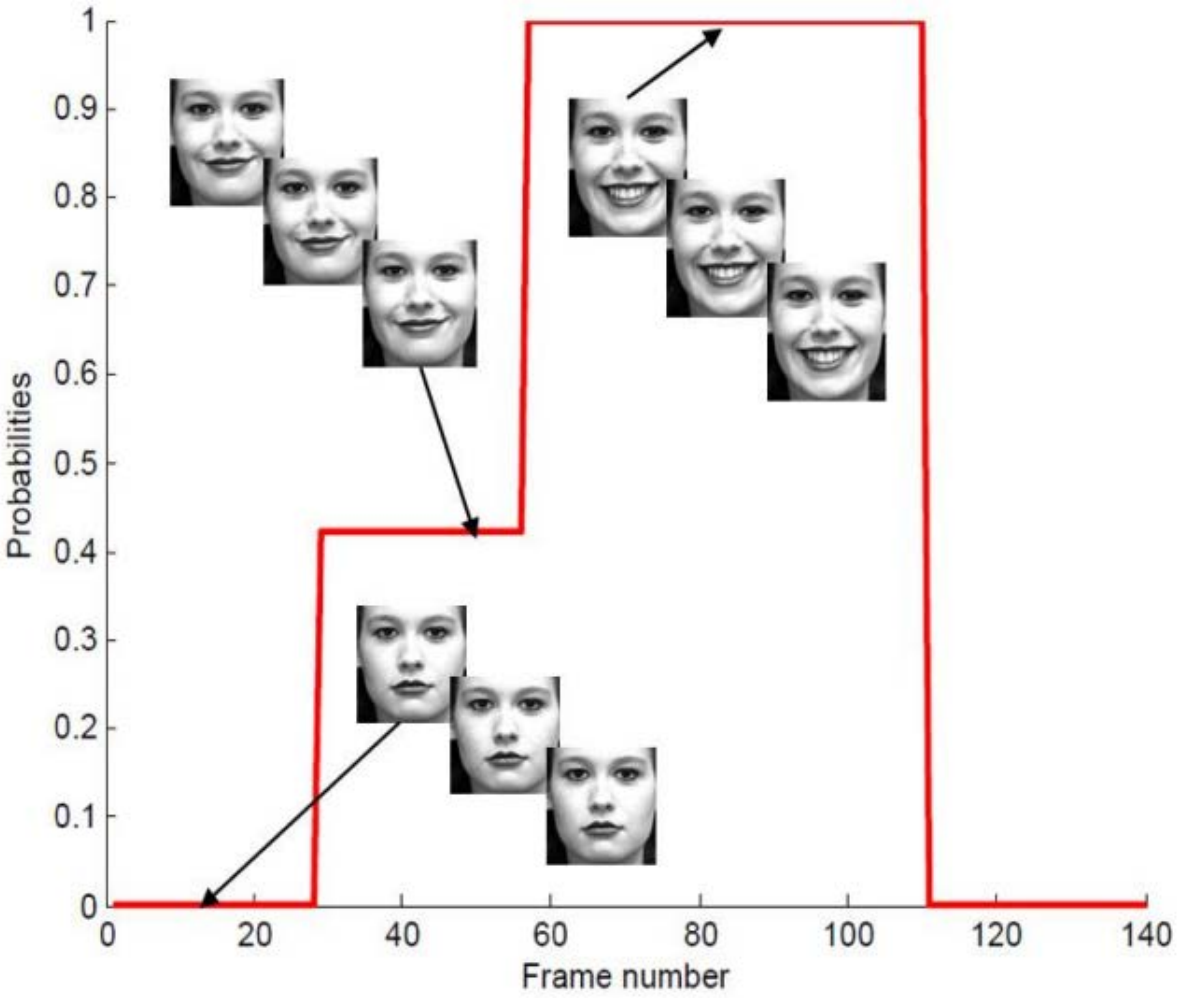
Comparison

Comparison with others on pain database

Algorithm	Accuracy	# subjects	# sequences
MIL-HMM	**85.23%**	25	149
MS-MIL (K. Sikka *(*et.al) [1])	83.7%	23	147
P. Lucey *et.al* [9]	80.99%	20	142
A.B. Ashraf *et.al* [10]	68.31%	20	142
A.B. Ashraf *et.al* [10]	81.21%	21	84

Segment Localization

Sample on CK+ Database



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!