A Novel Sequential Monte Carlo Framework For Predicting Ambiguous Emotion State

Jingyao Wu, Ting Dang, Vidhyasaharan Sethu, Eliathamby Ambikairajah

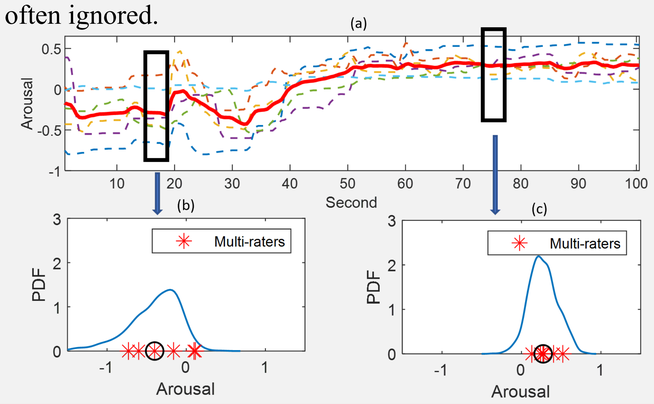
jingyao.wu@unsw.edu.au

1. Introduction

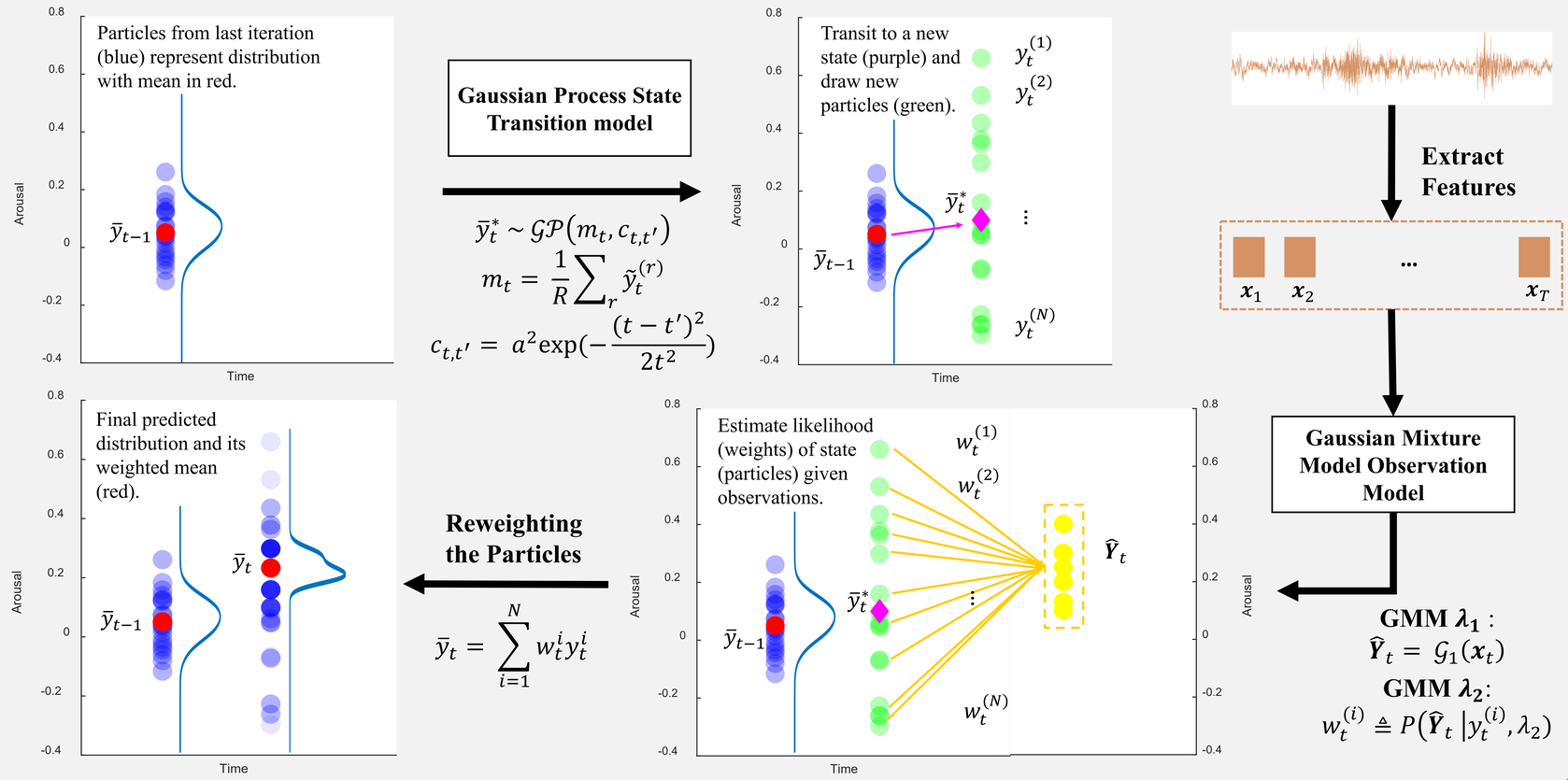
Emotion labels are typically obtained from multiple annotators. However, most automatic emotion recognition systems **ignore disagreement between annotators which reflects the ambiguity** and subtlety of emotions. In this paper, we **propose a novel Sequential Monte Carlo framework** that models the perceived emotion as time-varying distributions that allows for ambiguity to be incorporated.

Limitations of Current Systems

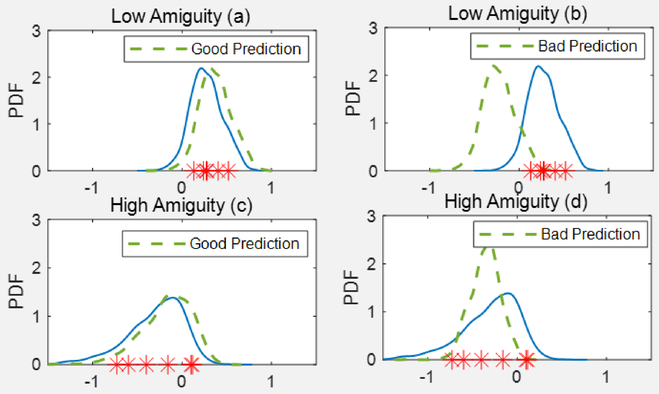
- Typically, only the **mean** rating (solid red line) is considered and inter-rater differences are ignored.
- When distributions are considered,
  - they are often assumed to be Gaussian.
  - **temporal dependencies** between consecutive predictions are often ignored.



2. Proposed Sequential Monte Carlo Framework



3. Proposed Measures



- Concordance correlation coefficient (CCC) between prediction standard deviation the standard deviation of the emotion ratings;
- Mean squared error (MSE) between prediction mean and ground truth mean label for different frames partitioned into deciles based on the standard deviation of the labels.

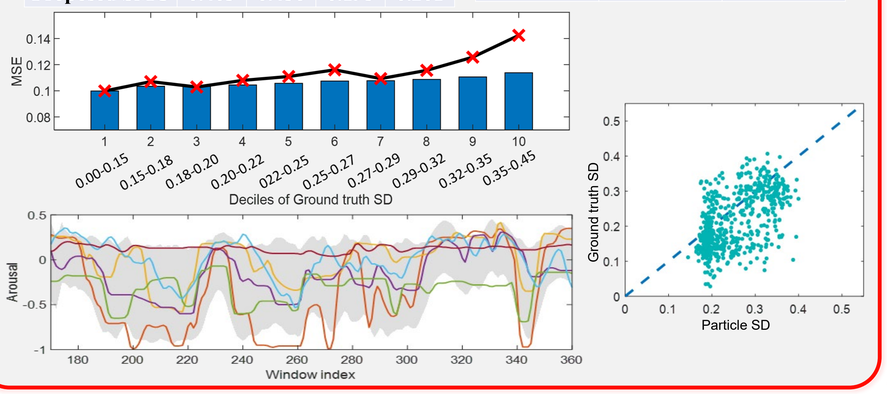
4. Results

Table 1 CCC and CC measure between predicted SD and SD from 6 annotators

	Arousal		Valence	
	CCC	CC	CCC	CC
BLSTM [17]	0.103	-	0.075	-
GMR [6]	-	0.568	-	0.132
Proposed SMC	**0.403**	**0.456**	**0.195**	**0.201**

Table 2 CCC measure between predicted mean and mean from 6 annotators

	Arousal	Valence
Literature [5]	0.783	0.495
Proposed SMC	**0.702**	**0.391**



5. Experimental Settings

- Corpus: the RECOLA dataset; 9 training & 9 development utterances.
- Arousal & valence labels; 6 annotators.
- 40ms sampling rate; 1 second window (50% overlap).
- Delay compensation: 4 seconds for arousal and 2 seconds for valence.
- 8 - mixture GMM for λ_1, 4 – mixture GMM for λ_2.
- 1000 particles.

6. Conclusion

- Inter-rater differences in emotion annotations reflecting ambiguity in the state can be represented as a series of non-parametric distributions.
- The proposed framework is able to track the level of ambiguity in the labels over time and predict the emotion state accurately within regions of low ambiguity.