**Верхоляк Оксана Владимировна Автоматическое распознавание эмоциональных состояний дикторов по голосовым характеристикам и тональности текста высказывания**

ОГЛАВЛЕНИЕ ДИССЕРТАЦИИ

кандидат наук Верхоляк Оксана Владимировна

Contents

Реферат

Synopsis

1 Introduction

1.1 Speech emotion recognition

1.2 Motivation for the current research

1.2.1 Dialogue modeling

1.2.2 Domain adaptation

1.2.3 Compactness vs. flexibility

1.2.4 Interpretable AI

1.2.5 Bimodal fusion

1.3 Practical application areas

1.4 Main goals of the thesis

1.5 Outline

2 Theoretical Concepts of Emotions

2.1 Definition of emotions

2.2 Emotion representation

2.2.1 Categorical approach

2.2.2 Dimensional approach

2.2.3 Mixed approach

2.3 Collection of emotional speech samples

2.3.1 Acted emotions

2.3.2 Induced emotions

2.3.3 Natural emotions

2.4 Annotation of emotions

2.4.1 Utterance-level

2.4.2 Frame-level

2.5 Challenges of emotion recognition

2.6 Summary

3 Background on Computational Methods

3.1 Feature extraction

3.1.1 Acoustic features

3.1.2 Linguistic features

3.1.3 Reducing feature space

3.2 Classification

3.2.1 Support vector machine

3.2.2 Logistic regression

3.2.3 Feed-forward neural networks

3.2.4 Recurrent neural networks

3.3 Generalization vs. overfitting

3.4 Performance evaluation

3.5 Summary

4 Data and Tools

4.1 Emotional speech corpora

4.1.1 IEMOCAP

4.1.2 RAMAS

4.1.3 CreativeIT

4.1.4 USoMS-e

4.2 Tonal dictionaries

4.2.1 SentiWordNet

4.2.2 SentiWS

4.3 Software tools

4.4 Summary

5 Acoustic Modeling

5.1 Motivation for the dialogue context modeling

5.2 Proposed system for dialogue-level context modeling

5.2.1 PCA-CCA-based domain adaptation

5.2.2 First-stage LSTM modeling

5.2.3 Data balancing

5.2.4 Second-stage LSTM modeling

5.3 Experimental setup

5.4 Experimental results

5.5 Discussion

5.6 Summary

6 Linguistic Modeling

6.1 Motivation for the proposed linguistic modeling

6.2 Proposed system for linguistic modelling

6.2.1 Machine translation

6.2.2 Extraction of polarity scores

6.2.3 Feature summarization

6.2.4 Feature selection

6.3 Experimental setup

6.3.1 Predefined data split

6.3.2 Cross-validation

6.3.3 Challenge trial submissions

6.4 Experimental results

6.4.1 SentiWordNet results

6.4.2 SentiWS results

6.4.3 Combination of SentiWordNet and SentiWS results

6.4.4 Comparative analysis

6.5 Discussion

6.6 Summary

7 Bimodal Emotion Recognition

7.1 Motivation for bimodal speech emotion recognition

7.1.1 Acoustic and linguistic features

7.1.2 Fusion strategies

7.1.3 Sentiments vs. emotions

7.2 Proposed method for bimodal speech emotion recognition

7.2.1 Acoustic modeling

7.2.2 Linguistic modeling

7.2.3 Bimodal fusion

7.3 Experimental setup

7.3.1 Annotation confidence vs. train data size

7.3.2 Bimodal speech emotion recognition experiments

7.4 Experimental results

7.4.1 Annotation confidence estimation results

7.4.2 Audio-based results

7.4.3 Text-based results

7.4.4 Bimodal fusion results

7.5 Discussion

7.6 Summary

8 Conclusion

8.1 Overall summary

8.2 Thesis contributions

8.2.1 Theoretical

8.2.2 Practical

8.2.3 Experimental

8.3 Limitations and future research

Appendix

References

Acronyms

List of Figures

List of Tables

List of Own Publications

Patents