Affect Representation Schemes used in Affective Video Content Analysis

A SYSTEMATIC REVIEW

Author : Ashika Chakravorty (AshikaChakravorty@student.tudelft.nl) **Supervisor** : Bernd Dudzik **Responsible Professor** : Chirag Raman

1. Introduction

Background:

- Affective content of a video refers to the emotions, mood, or feelings conveyed in the video and experienced by viewers [1]
- Affect: broad term for a range of subjective experience
- Affective Video Content Analysis (AVCA) aims to automatically recognize the affective content of videos [2]
- Perspectives of AVCA:
  - Affective Movie Content Analysis (AMCA): analyses emotions within video stimuli
  - Video Emotion Recognition (VER): estimates expressed emotion of agent in a video, often induced by a stimulus
- Approaches of AVCA:
  - Direct analysis: infers affective content from video's audiovisual features
  - Implicit analysis: recognizes affective content elicited by videos by using users' visible behaviour or physiological signals
- Affect is represented by Affect Representation Schemes (ARS), which can be categorical (Fig. 1) or dimensional (Fig. 2)

Fig 1. Ekman's basic emotions - anger, fear, disgust, surprise, happiness, sadness [3]

Fig 2. Russell's model - Valence & Arousal (VA) [4]

Motivation:

- Significant variation in use of ARS
- Lack of consensus among experts on ARS choice for AVCA
- Accuracy and effectiveness of AVCA systems can be influenced by selected ARS

References

[1] J. Tarvainen, M. Sjoberg, S. Westman, J. Laaksonen, and P. Oittinen, "Content-based prediction of movie style, aesthetics, and affect: Data set and baseline experiments," IEEE Transactions on Multimedia, vol. 16, pp. 2085–2098, 8 2014.
[2] Y. Baveye, C. Chamaret, E. Dellandrea, and L. Chen, "Affective video content analysis: A multidisciplinary insight," IEEE Transactions on Affective Computing, vol. 9, pp. 396–409, 4 2018.
[3] P. Ekman, "Basic emotions," in Handbook of Cognition and Emotion. John Wiley Sons, Ltd, 1999, ch. 3, pp. 45–60.
[4] J. Russell, "A circumplex model of affect," Journal of Personality and Social Psychology, vol. 39, pp. 1161–1178, 1980.
[5] M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., "The prisma 2020 statement: An updated guideline for reporting systematic reviews," BMJ, vol. 372, 2021.
[6] J. Panksepp, Affective neuroscience: The foundations of human and Animal Emotions. Oxford University Press, 1998.

2. Research Question

"How are various affect representation schemes currently used in Affective Video Content Analysis?"

Sub-questions:

- What different types of input data are used?
- What types of affective states are targeted?
- What different ARS are used?
- Are ARS used in combination or individually?
- Are used ARS based on psychological theory?
- Are there differences in ARS used in direct & implicit analysis?
- What is the popularity of types ARS & over time?

3. Methodology

- Conducted a systematic literature review (SLR) following PRISMA guidelines [5].
- Conducted search on 3 databases - IEEE Explore, Scopus, Web of Science

Eligibility criteria:

- published after 2008 in English
- original *research papers* published in journals or conference proceedings
- limited to the field of Computer Science
- paper is based on AVCA: perspective is AMCA and not VER
- paper includes the use of ARS

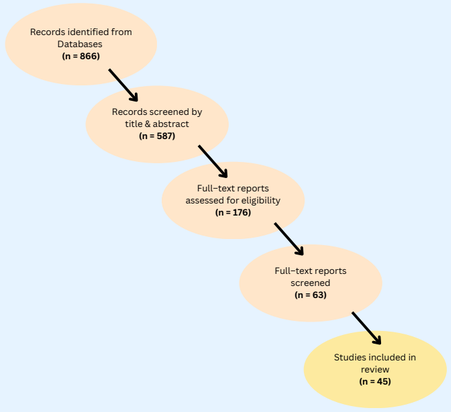

Fig. 3 PRISMA flow diagram of search results [5]

4. Results

1.Types of ARS used in AVCA:

- Dimensional ARS: Russell's Valence-Arousal (VA) model used
- Categorical ARS: distinct discrete categorizations of emotions, often including the neutral affect state
- Combinational ARS: categorize dimensional models into quadrants
- Only 28% of papers provided motivations for their choice of ARS

2.Targeted Affective States:

- 90% of papers focused solely on emotions, while 10% also considered mood, attitude, and violence alongside emotions

3.Basis of ARS:

- Only 46% of papers mentioned the basis for their chosen ARS
- The majority of ARS were based on psychological theories (Fig. 5), with one exception based on affective neuroscience research [6]

4.Input Data used in ARS:

- Direct analysis: audiovisual data and text from movie scripts
- Implicit analysis: physiological signals, visible behaviours, user annotation and text data from user movie reviews or comments
- Direct & implicit: combinations of input data
- Distribution can be seen in Fig. 7
- No significant relationship was found between input data and types ARS used
- Most papers utilized unique input data based on system design and requirements

5.Popularity of ARS:

- Dimensional and categorical approaches equally popular (Fig. 4)
- Russell's VA model was the most preferred (Fig. 5)
- Popularity of ARS types fluctuated over time, but no significant trend was found (Fig. 6)

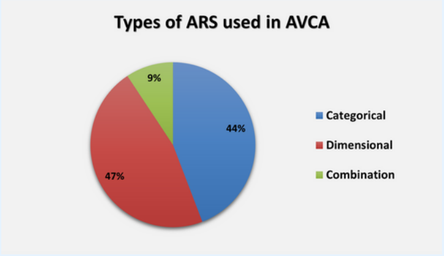

Fig. 4 Popularity of types of ARS

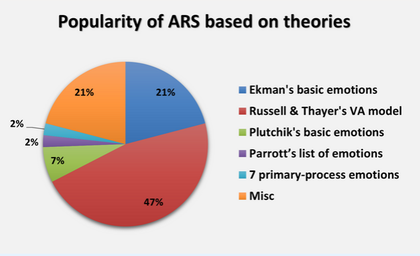

Fig. 5 Distribution of ARS based on psychological theories

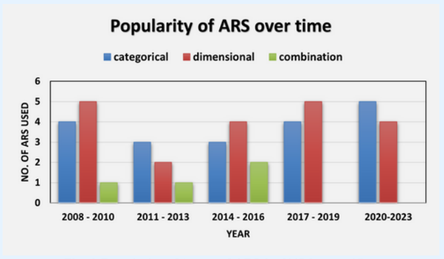

Fig. 6 Popularity of types of ARS in AVCA over 15 years

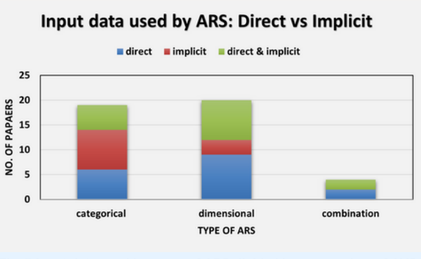

Fig. 7 Input data used by types of ARS divided by approaches of AVCA

5. Discussion

- Researchers should provide motivations and enhance transparency in reporting their choice of ARS to improve reliability of systems
- High preference for Russell's VA model suggests recognition of its advantages in capturing complex emotional experiences
- Categorical ARS: Varying terminologies used to describe similar emotions, sometimes leading to misinterpretation between emotion and mood

Limitations:

- Time constraints led to screening of a limited number of papers (63 out of 176), limiting data available for comprehensive analysis
- There could be potential bias introduced as review was performed by a single researcher
- Limited similarities found between input data and types of ARS, possibly due to insufficient data
- Lack of significant trends in the popularity of ARS may be due to limited data

6. Conclusion

- Wide range of ARS and input data used: indicates the dynamic nature of AVCA research
- Dimensional ARS is popular, but limited data prevents conclusive evidence

Future work:

- Further research including more papers needed to draw concrete conclusions in AVCA
- Evaluation of ARS used in VER should be explored
- The impact of different ARS on the performance of AVCA systems should be investigated

TU Delft