



OCEAN: Towards Developing an Opportunistic Continuous Emotion Annotation Framework

<u>Akhilesh Adithya</u>¹, Snigdha Tiwari¹, Sougata Sen¹, Sandip Chakraborty², Surjya Ghosh¹

1 - Department of Computer Science and Information Systems, BITS Pilani Goa, INDIA 2 - Department of Computer Science and Engineering, IIT Kharagpur, INDIA



Typical ML-based emotion detection model

- Emotion detection \rightarrow ML model
 - Combination of different modalities and emotion ground truth
- Emotion ground truth → typically collected as <u>manual self-report</u>



- Post interaction/ post stimuli
 - Issue fails to capture emotion wrt intra-video subtle nuances



- Post interaction/ post stimuli
 - Issue fails to capture emotion wrt intra-video subtle nuances



- Continuous
 - Issues degrades viewing experience ; immersion ; adds strain



- Post interaction/ post stimuli
 - Issue fails to capture emotion wrt intra-video subtle nuances



- Continuous
 - Issues degrades viewing experience ; immersion ; adds strain

Need for an annotation approach to capture emotion with respect to intra-video nuances without (or minimal) disruption of the viewing experience

Problem Statement

Develop a new framework for emotion annotation that captures the intra video nuances, and at the same time reduce the number of probes

Dataset : CASE (Continuously Annotated Signals of Emotion)^[1]



Image of CASE Dataset Setup

A continuous dataset that annotates valence and arousal along with the outputs of 6 sensors.

- Electrocardiograph (ECG)
- Blood Volume Pulse (BVP)
- Galvanic Skin Response (GSR)

- Respiration (RSP)
- Skin Temperature (SKT)
- Electromyography (EMG)

[1] - Sharma, K., Castellini, C., van den Broek, E.L. *et al.* A dataset of continuous affect annotations and physiological signals for emotion analysis. *Sci Data* **6**, 196 (2019). <u>https://doi.org/10.1038/s41597-019-0209-0</u>

Dataset



Video id	Emotion	Valence	Arousal	Duration (in sec.)		
1	amusing	med/high	med/high	185		
2	amusing	med/high	med/high	173		
3	boring	low	low	119		
4	boring	low	low	160		
5	relaxing	med/high	low	145		
6	relaxing	med/high	low	147		
7	scary	low	high	197		
8	scary	low	high	144		

CASE dataset video details

- 30 users
- 8 videos in randomised order
- 2D plane Joystick Input
- Emotional data collected -
 - Valence & Arousal : Scale 1-9

Key Intuition behind the proposed approach

- Emotions change slowly as they have a residual effect.
- By exploring the dataset, we can see that these changes are reflected in physiological signals.
- Measuring these changes would allow us to probe at an opportune point.
- These probes would allow us to gather the intra-video nuances with minimal number of disturbances.

In summary, opportunistic annotation instead of continuous annotation

Take Physiological Signals





















Results

What is the reduction in the number of probes?



Baseline probe - Every
5 seconds in continuous
mode of annotation

We see that there is an average decrease in probes by **89%**

Does reducing the number of probes affect annotation quality?



Change of arousal and valence against time

Does reducing the number of probes affect annotation quality?

Video id	Valence					Arousal						
	Min (Act.)	Min (Sam.)	Max (Act.)	Max (Sam.)	Med (Act.)	Med (Sam.)	Min (Act.)	Min (Sam.)	Max (Act.)	Max (Sam.)	Med (Act.)	Med (Sam.)
1	1.763	1.883	9.5	9.5	6.266	6.284	3.164	3.552	9.013	9.013	5.248	5.624
2	3.058	3.514	9.5	9.5	6.169	6.562	1.71	2.936	8.414	8.396	5.008	5.101
3	2.42	2.967	7.39	7.39	5	5	0.509	0.509	7.005	6.375	3.772	4.844
4	2.686	3.145	7.943	7.572	5.009	5.05	0.5	1.402	6.477	6.477	3.836	4.065
5	1.163	1.8	9.5	9.5	6.052	5.845	0.5	0.5	6.291	5.937	4.115	4.337
6	2.717	2.981	9.206	7.705	5.397	5.702	2.905	2.912	6.83	6.609	5	5
7	0.5	0.5	7.813	7.513	2.887	3.2	1.023	1.023	9.5	9.5	7.200	7.711
8	0.5	0.5	8.919	6.871	4.613	3.198	3.018	4.287	9.5	9.5	6.45	7.417

Table : Minimum, median and maximum values of arousal and valence for Actual and Sampled values

Does reducing the number of probes affect annotation quality?

Video id	Valence					Arousal						
viaco iu	Min (net.)	Min (Sam)	Max (men)	Max (Sam.)	Micu (Act.)	Med (Sam)	Nim (men)	Min (Sam.)	Max (rich)	Max (Sam.)	Med (men)	Med (Sam)
1	1.763	1.883	9.5	9.5	6.266	6.284	3.164	3.552	9.013	9.013	5.248	5.624
2	2.059	2 5 1 /	0.5	0.5	6.160	6.562	1.71	2.026	9 / 1 /	<u> 206</u>	5.009	5 101
3	2.42	2.967	7.39	7.39	5	5	0.509	0.509	7.005	6.375	3.772	4.844
4	2.686	3.145	7.943	7.572	5.009	5.05	0.5	1.402	6.477	6.477	3.836	4.065
5	1.163	1.8	9.5	9.5	6.052	5.845	0.5	0.5	6.291	5.937	4.115	4.337
6	2.717	2.981	9.206	7.705	5.397	5.702	2.905	2.912	6.83	6.609	5	5
7	0.5	0.5	7.813	7.513	2.887	3.2	1.023	1.023	9.5	9.5	7.200	7.711
8	0.5	0.5	8.919	6.871	4.613	3.198	3.018	4.287	9.5	9.5	6.45	7.417

Table : Minimum, median and maximum values of arousal and valence for Actual and Sampled values

What is the number of probes for the CASE dataset?



We see that the average number of probes is 4 per user per video

Conclusions

- A novel framework designed for emotion annotation, OCEAN detects significant changes in the input signals and decides whether or not to probe for information using k-means clustering of change point scores
- Reduced number of probes by 89% on average, yet captures similar annotation scores for valence and arousal

Conclusions

- A novel framework designed for emotion annotation, OCEAN detects significant changes in the input signals and decides whether or not to probe for information using k-means clustering of change point scores
- Reduced number of probes by 89% on average, yet captures similar annotation scores for valence and arousal

Future Work

- Create a working prototype and run live user studies
- Quantify storage reduction and effects of latency of probing window size
- Generalizability of the framework (validation on other datasets)

Thank You



Contact Me: f20190044@goa.bits-pilani.ac.in



Website: https://akhileshadithya.github.io



Github: https://github.com/AkhileshAdithya