

# Change Point Analysis (CPA) of the CASE Dataset

Lucy Roellecke

22 November 2023

## Abstract

This document summarizes the results of performing a change point analysis (CPA) of the textitDataset of continuous affect annotations and physiological signals for emotion analysis (CASE, available here: [https://springernature.figshare.com/collections/A\\_dataset\\_of\\_continuous\\_affect\\_annotations\\_and\\_physiological\\_signals\\_for\\_emotion\\_analysis/4260668](https://springernature.figshare.com/collections/A_dataset_of_continuous_affect_annotations_and_physiological_signals_for_emotion_analysis/4260668)). The results reported in this document are in line with the results found by the original authors. Qualitatively meaningful change points in the participants' annotation data can be detected using a data-driven approach such as the CPA easily for highly arousing, negative videos (such as scary videos) but are harder to detect in low arousing and positively valenced videos. Interindividual differences are high in all conditions.

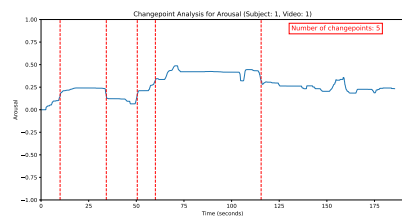
## 1 Analysis Notes

The change point analysis was undertaken by using the Python package **ruptures** (Version 1.1.8) and the algorithm of linearly penalized segmentation (**Pelt**). The parameters of the algorithm were set as follows: **model=12** (least squared deviation model), **jump=5** (controls the grid of possible change points), and **pen=1** (the optimal penalization factor was determined by visual inspection of an elbow plot with possible penalty values on the x-axis and the resulting number of change points on the y-axis). **cpa.py** performs a change point analysis for each participant separately, while **cpa\_average.py** performs a change point analysis for the annotation data averaged across all participants. Both scripts return csv files with the x-(time)-coordinates of the change points and the number of change points, and plot the time-series together with the change points as dashed lines for each of the 8 stimuli videos separately (see 2). The latter script performs different kinds of visualization: 1) Plotting the averaged timeseries and the averaged change points; 2) Plotting all timeseries of all participants and the averaged change points; 3) Plotting the averaged timeseries and all change points for all participants. See 2 for the different visualizations. For code and figure availability, see section 3.

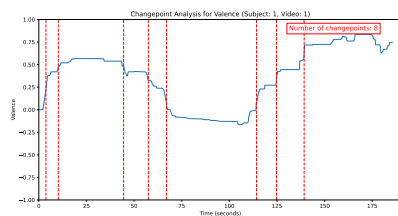
## 2 Descriptive Results

### 2.1 CPA for each participant separately

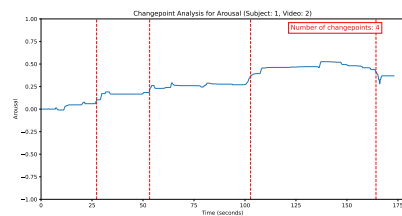
Figure 1 shows the change points identified for each video’s arousal and valence ratings for subject 1 (as an example). The algorithm seems to be able to identify abrupt decreases and increase in the emotion ratings.



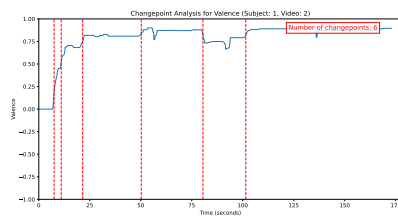
(a)



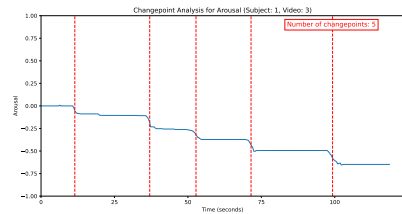
(b)



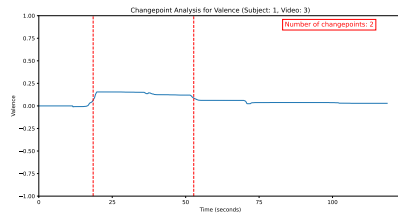
(c)



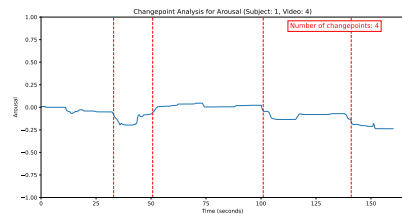
(d)



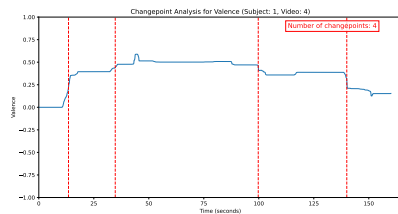
(e)



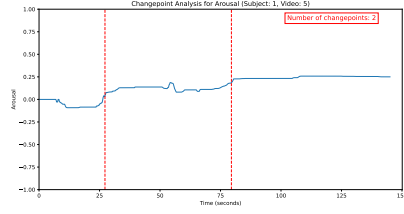
(f)



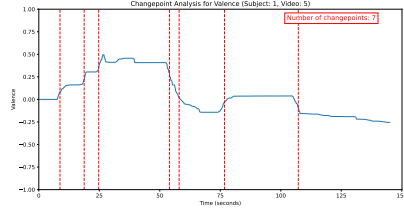
(g)



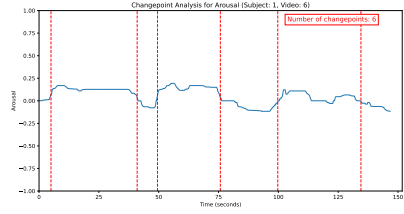
(h)



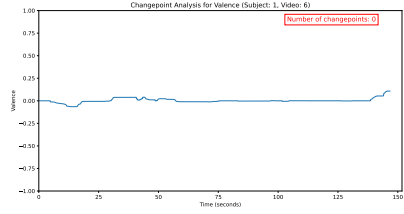
(i)



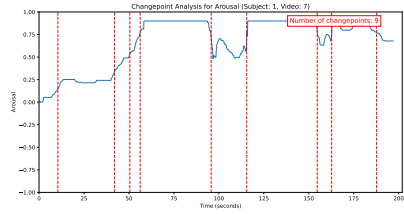
(j)



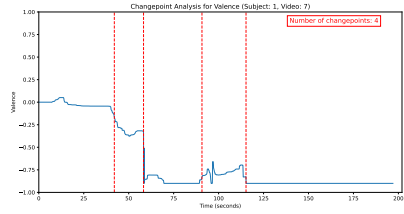
(k)



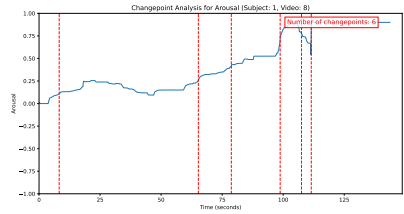
(l)



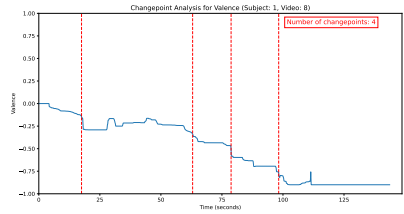
(m)



(n)



(o)



(p)

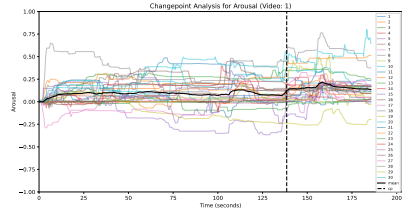
Figure 1: *Change points in arousal and valence ratings for all videos for one subject.* Change points as identified by the algorithm for each video's arousal (left column) and valence (right column) ratings for subject 1. Red dashed lines indicate the change points, the thin blue line is the subject's continuous rating. The number of change points identified is given in the top right corner of each plot. Note the different lengths of the videos.

Table 1: *X-Coordinates (timestamps) of change points for all videos averaged across participants.* Q = Quadrant (video category); HP = high arousal and positive valence; LN = low arousal and negative valence; LP = low arousal and positive valence; HN = high arousal and negative valence; CP = change point; No. = number.

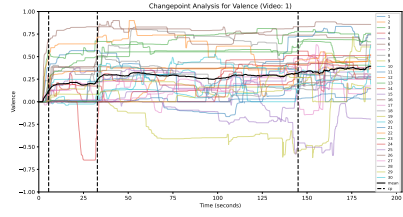
Video	Q	CP Valence	No. CP Valence	CP Arousal	No. CP Arousal	Model	Jump	Penalty
1	HP	[5.5, 32.75, 145.0]	3	[138.5]	1	l2	5	1
2	HP	[13.75, 61.25, 116.5, 141.25]	4	[117.75]	1	l2	5	1
3	LN	[45.0]	1	[19.75, 42.5]	2	l2	5	1
4	LN	[]	0	[20.5, 68.25]	2	l2	5	1
5	LP	[7.0, 101.5]	2	[37.75, 59.75, 95.5]	3	l2	5	1
6	LP	[16.75]	1	[]	0	l2	5	1
7	HN	[42.75, 59.0, 71.5, 90.5, 116.0]	5	[43.5, 59.0, 90.5, 115.5, 150.75]	5	l2	5	1
8	HN	[61.0, 76.75, 101.25, 112.0]	4	[18.0, 57.25, 73.25, 99.75, 111.75]	5	l2	5	1

## 2.2 CPA for data averaged across participants

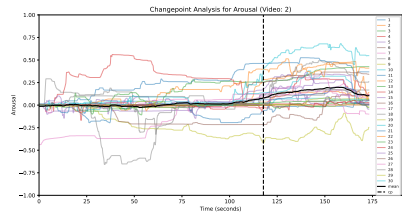
Figure 2 shows the change points identified for each video’s arousal and valence ratings averaged across subjects. Additionally, all subjects’ annotation timeseries for valence and arousal are depicted to show the interindividual differences in ratings. Even though the algorithm seems to be able to identify change points in all videos, only video 7 and video 8 seemed to have significant changes in valence and arousal ratings. As those were the two videos labelled as ‘scary’ by the authors, it seems plausible that participants’ rating varied more in these videos and were more extreme than in the other videos. For the other two high arousing but positive videos (1 and 2), more change points have been identified than in the low arousing videos, albeit this effect is less than in the negative videos. Note the high differences between individual participants in their emotion ratings. See table 1 for the x-coordinates (timestamps) of the averaged change points for all videos and for both arousal and valence.



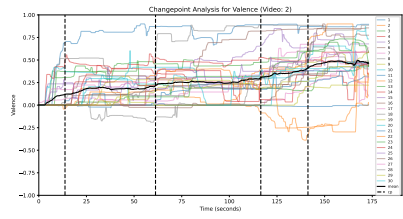
(a)



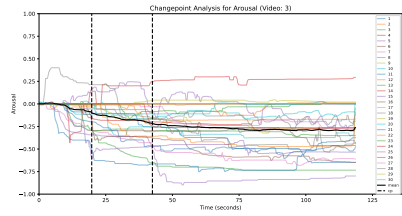
(b)



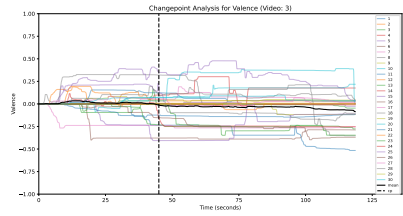
(c)



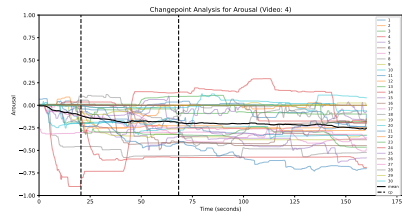
(d)



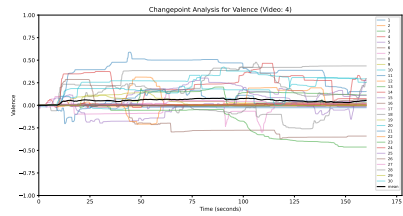
(e)



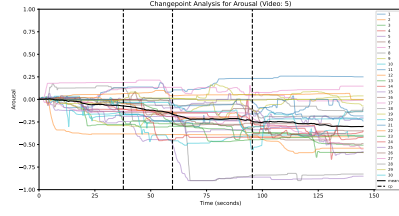
(f)



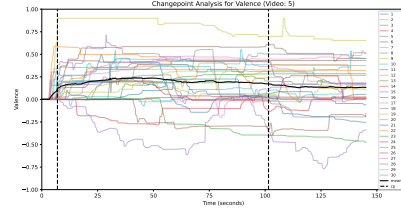
(g)



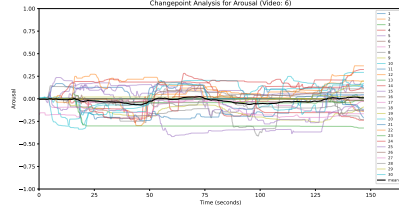
(h)



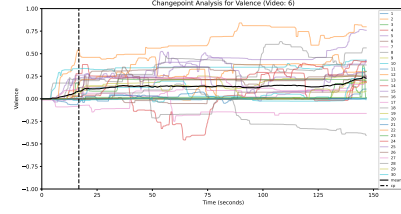
(i)



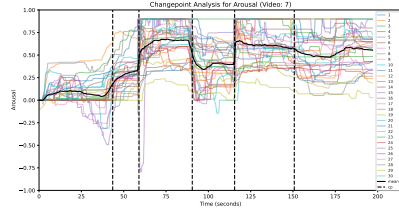
(j)



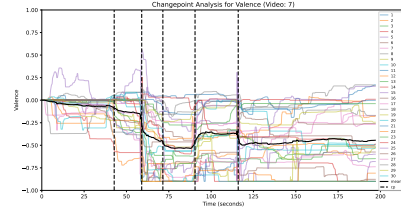
(k)



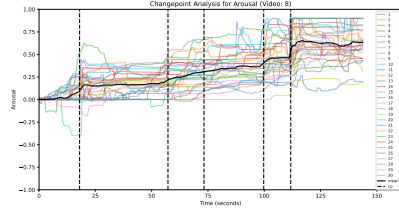
(l)



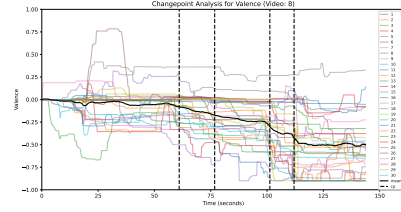
(m)



(n)



(o)



(p)

Figure 2: *Change points in arousal and valence ratings for all videos averaged across subjects* Change points as identified by the algorithm for each video's arousal (left column) and valence (right column) ratings averaged across subjects. Black dashed lines indicate the change points. Colored lines indicate the ratings of the different participants, the thick black line is the mean rating. Even though the algorithm seems to be able to identify change points in all videos, only video 7 and video 8 (high arousal and negative valence) seemed to have significant changes in valence and arousal ratings. Note the different lengths of the videos.

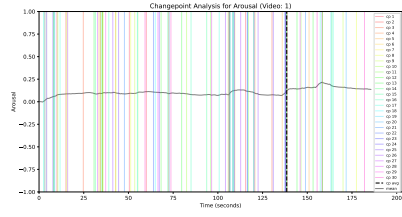
### 2.3 CPA for data averaged across participants but with individual change points

Figure 3 shows the change points identified for each video's arousal and valence ratings for all subjects in one plot to show the interindividual differences in change points. Additionally, the mean timeseries for valence and arousal and the mean change points for all participants are depicted. We can see that there are high differences both in the number and in the point in time of change points between participants. Only in video 7 and 8, there seems to be a bigger congruence of participants' data. As already mentioned, these videos were the ones labelled as 'scary' by the authors. One explanation for this pattern is that these scary videos succeeded in eliciting more similar emotional reactions across participants than the positive and low arousing videos where participants differed more in both timing and magnitude of changes in annotation data.

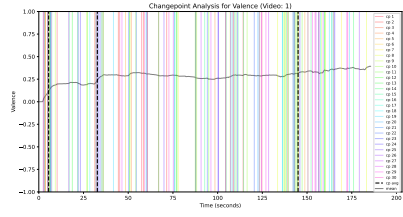
## 3 Code and Figure Availability

All python scripts can be found on GitHub: <https://github.com/lucyroe/AVR.CPA>. All figures depicted here can also be found separately on OwnCloud in the AffectiveVR directory.

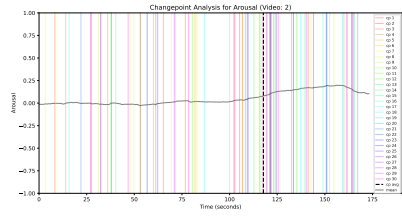




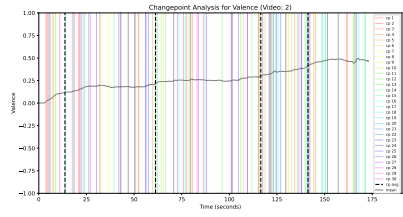
(a)



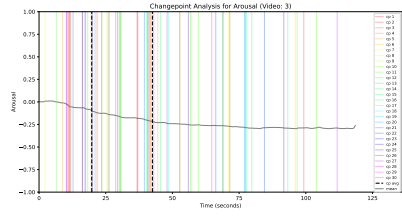
(b)



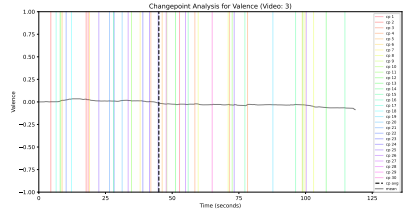
(c)



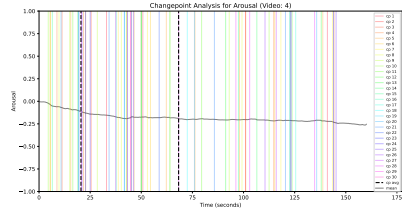
(d)



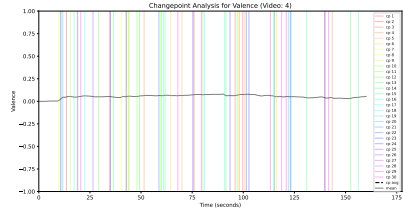
(e)



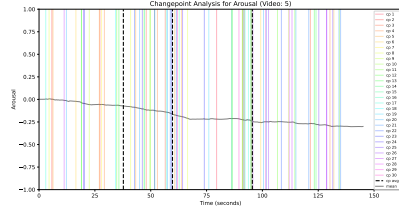
(f)



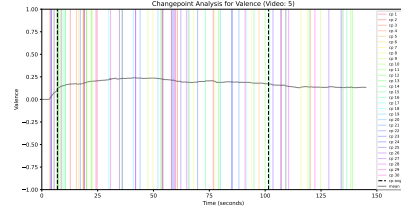
(g)



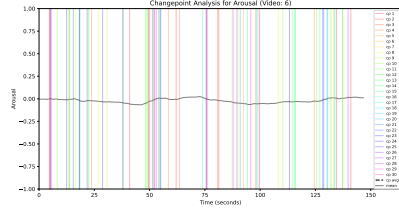
(h)



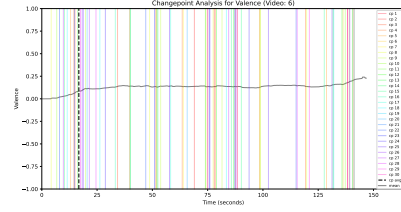
(i)



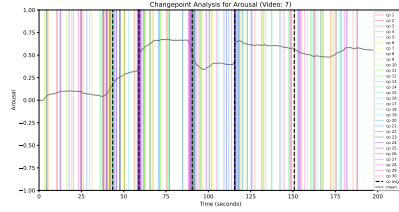
(j)



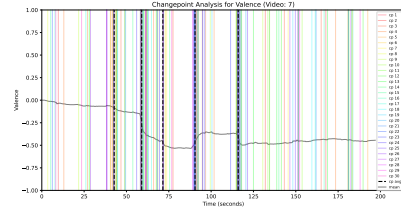
(k)



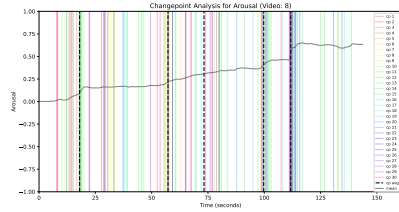
(l)



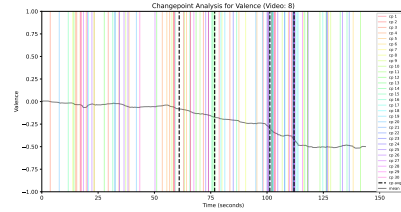
(m)



(n)



(o)



(p)

Figure 3: *Change points in arousal and valence ratings for all videos for all subjects* Change points as identified by the algorithm for each video's arousal (left column) and valence (right column) ratings for all subjects. Thin colored lines indicate the subjects' change points. Black dashed lines indicate the mean change points. The thin grey line is the mean time series. There seem to be great interindividual differences. Only in video 7 and video 8, participants' alignment seemed to be higher. Note the different lengths of the videos.