Paper：《EmotionMeter: A Multimodal Framework for Recognizing Human Emotions》

（https://bcmi.sjtu.edu.cn/~lubaoliang/papers/2019/1.pdf）

理论基础：

To elicit specific emotions in the experimental environment, we used carefully selected film clips as the stimuli. The reliability of film clips with audiovisual stimuli in eliciting emotions has been studied in the literature [59]–[61].

[59] J. J. Gross and R. W. Levenson, “Emotion elicitation using films,” Cogn. Emotion, vol. 9, no. 1, pp. 87–108, 1995.

[60] A. Schaefer, F. Nils, X. Sanchez, and P. Philippot, “Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers,” Cogn. Emotion, vol. 24, no. 7, pp. 1153–1172, 2010.

[61] Y. Baveye, E. Dellandréa, C. Chamaret, and L. Chen, “LIRIS-ACCEDE: A video database for affective content analysis,” IEEE Trans. Affect. Comput., vol. 6, no. 1, pp. 43–55, Jan./Mar. 2015.

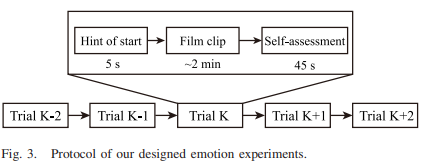
视频剪辑要求：

The criteria for clip selection were as follows: 1) the length of the videos should not be too long to cause visual fatigue; 2) the videos should be understood without explanation; and 3) the videos should elicit a single desired target emotion.

实验设计：

There were 168 film clips in total for four emotions (happy, sad, fear, and neutral) in our material pool, and forty-four participants (22 females, all college students) were asked to assess their emotions when watching the film clips with keywords of emotions (happy, sad, neutral, and fear) and ratings out of ten points (from -5 to 5) for two dimensions: 1) valence and 2) arousal. The valence scale ranges from sad to happy.

The duration of each film clip was approximately two minutes. Fig. 3 presents the detailed protocol of our designed emotion experiments. Each film clip had a 5 s hint for starting and a 45 s self-assessment with the PANAS scales [62] after each clip.



[62] D. Watson, L. A. Clark, and A. Tellegen, “Development and validation of brief measures of positive and negative affect: The PANAS scales,” J. Pers. Soc. Psychol., vol. 54, no. 6, pp. 1063–1070, 1988.

Paper：《Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks》

（https://bcmi.sjtu.edu.cn/~lubaoliang/papers/2015/2015-16.pdf）

理论基础：

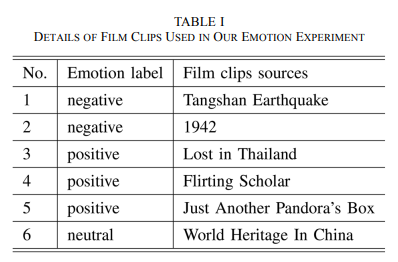
Nowadays, there are various kinds of stimuli used in emotion research like image, music, metal imagery, and films. Compared to other stimuli, emotional films have several advantages. The existing studies have already evaluated the reliability and efficiency of film clips to elicitation [50], [51]. Emotional films contain both scene and audio, which can expose subjects to more real-life scenarios and elicit strong subjective and physiological changes.

[50] J. J. Gross and R. W. Levenson, “Emotion elicitation using films,” Cogn. Emotion, vol. 9, no. 1, pp. 87–108, 1995.

[51] A. Schaefer, F. Nils, X. Sanchez, and P. Philippot, “Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers,” Cogn. Emotion, vol. 24, no. 7, pp. 1153–1172, 2010.

实验设计：

There are totally fifteen clips in one experiment and each of them lasts for about 4 min.



评估：

We selected the subjects using the Eysenck Personality Questionnaire (EPQ). The EPQ is a questionnaire to assess the personality traits of a person devised by Eysenck et al.[52].

[52] S. B. Eysenck, H. J. Eysenck, and P. Barrett, “A revised version of the psychoticism scale,” Pers. Individ. Differences, vol. 6, no. 1, pp. 21–29, 1985

Paper：《Comparing Recognition Performance and Robustness of Multimodal Deep Learning Models for Multimodal Emotion Recognition》

（https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9395500）

使用不同信号进行情绪识别：

Traditional emotion recognition systems are built with non-physiological signals [4], [5]. However, emotions also contain reactions from the central and peripheral nervous systems. Besides, electroencephalography (EEG)-based emotion recognition has been demonstrated to be a reliable method because of its high recognition accuracy, objective evaluation and stable neural patterns [6], [7], [8], [9], [10].

Moreover, other physiological signals such as electromyogram, electrocardiogram, skin conductivity, respiration, and eye movement signals are also used to recognize emotions [18], [19].

[4] B. Ko, “A brief review of facial emotion recognition based on visual information,” Sensors, vol. 18, no. 2, p. 401, 2018.

[5] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, “Current state of text sentiment analysis from opinion to emotion mining,” ACM Computing Surveys (CSUR), vol. 50, no. 2, p. 25, 2017.

[6] W.-L. Zheng and B.-L. Lu, “Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks,” IEEE Transactions on Autonomous Mental Development, vol. 7, no. 3, pp. 162–175, 2015.

[7] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, “Identifying stable patterns over time for emotion recognition from eeg,” IEEE Transactions on Affective Computing, vol. 10, no. 3, pp. 417–429, 2019.

[8] Y. Yang, Q. J. Wu, W.-L. Zheng, and B.-L. Lu, “EEG-based emotion recognition using hierarchical network with subnetwork nodes,” IEEE Transactions on Cognitive and Developmental Systems, vol. 10, no. 2, pp. 408–419, 2018.

[9] Z. Yin, Y. Wang, L. Liu, W. Zhang, and J. Zhang, “Cross-subject EEG feature selection for emotion recognition using transfer recursive feature elimination,” Frontiers in Neurorobotics, vol. 11, p. 19, 2017.

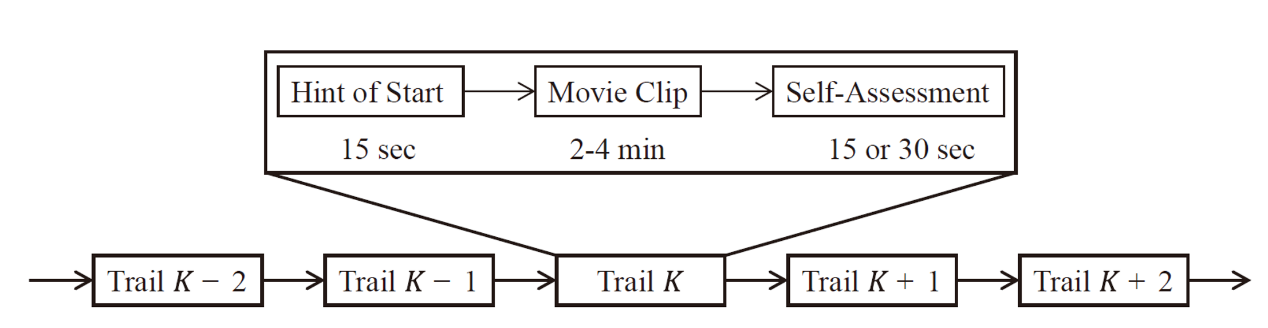
[10] R. Fourati, B. Ammar, J. Sanchez-Medina, and A. M. Alimi, “Unsupervised learning in reservoir computing for EEG-based emotion recognition,” IEEE Transactions on Affective Computing, pp. 1–1, 2020.

[18] J. Kim and E. Andre, “Emotion recognition based on physiological ´ changes in music listening,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, pp. 2067–2083, 2008.

[19] M. L.-H. Vo, A. M. Jacobs, L. Kuchinke, M. Hofmann, M. Conrad, ˜ A. Schacht, and F. Hutzler, “The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect,” Psychophysiology, vol. 45, no. 1, pp. 130–140, 2008

实验设计：（https://bcmi.sjtu.edu.cn/home/seed/seed-v.html）

All stimulus materials have 15 seconds before playing to introduce the background of the materials and the emotions they want to induce. After the stimulus materials are played, there will be 15 seconds or 30 seconds of self-evaluation and rest time according to the type of material. If the type of stimulus material was disgust or fear, the rest time was 30 seconds, and the time for happiness, neutrality and sadness was 15 seconds. The specific process is shown in the figure below.



Paper：《A Multimodal Database for Affect Recognition and Implicit Tagging》

（https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5975141）

理论基础：

Although the most straightforward way to represent an emotion is to use discrete labels such as fear or joy, label-based representations have some disadvantages. Specifically, labels are not cross-lingual: Emotions do not have exact translations in different languages, e.g., “disgust” does not have an exact translation in Polish [12]. Psychologists therefore often represent emotions or feelings in an n-dimensional space (generally 2 or 3D). The most famous such space, which is used in the present study and originates from cognitive theory, is the 3D valence, arousal-dominance or pleasure-arousal-dominance (PAD) space [13].

[12] J.A. Russell, “Culture and the Categorization of Emotions,” Psychological Bull., vol. 110, no. 3, pp. 426-450, 1991.

[13] J.A. Russell and A. Mehrabian, “Evidence for a Three-Factor Theory of Emotions,” J. Research in Personality, vol. 11, no. 3, pp. 273-294, Sept. 1977.

剪辑视频来源：

In a preliminary study, 155 video clips containing movie scenes manually selected from 21 commercially produced movies were shown to more than 50 participants; each video clip received 10 annotations on average [15]. The preliminary study was conducted utilizing an online affective annotation system in which the participants reported their emotions in response to the videos played by a web-based video player.

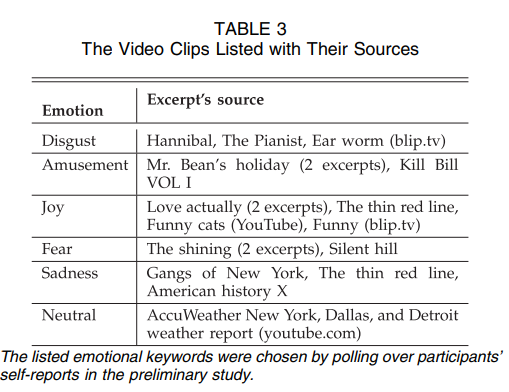
Fourteen video clips were chosen based on the preliminary study from the clips which received the highest number of tags in different emotion classes, e.g., the clip with the highest number of sad tags was selected to induce sadness. Three other popular video clips from online resources were added to this set (two for joy and one for disgust). Three past weather forecast reports (retrieved from youtube.com) were also used as neutral emotion clips. The videos from online resources were added to the data set to enable us to distribute some of the emotional video samples with the multimodal database described below. The full list of videos is given in Table 3.

Ultimately, 20 videos were selected to be shown which were between 34.9 and 117 s long (M = 81:4 s; SD = 22:5 s). Psychologists recommended videos from 1 to 10 minutes long for elicitation of a single emotion [17], [18]. Here, the video clips were kept as short as possible to avoid multiple emotions or habituation to the stimuli while keeping them long enough to observe the effect.

[15] M. Soleymani, J. Davis, and T. Pun, “A Collaborative Personalized Affective Video Retrieval System,” Proc. Third Int’l Conf. Affective Computing and Intelligent Interaction and Workshops, Sept. 2009.

[17] A. Schaefer, F. Nils, X. Sanchez, and P. Philippot, “Assessing the Effectiveness of a Large Database of Emotion-Eliciting Films: A New Tool for Emotion Researchers,” Cognition and Emotion, vol. 24, no. 7, pp. 1153-1172, 2010.

[18] J. Rottenberg, R.D. Ray, and J.J. Gross, “Emotion Elicitation Using Films,” Handbook of Emotion Elicitation and Assessment, series in affective science, pp. 9-28, Oxford Univ. Press, 2007.



Paper：《DECAF: MEG-Based Multimodal Database for Decoding Affective Physiological Responses》

（https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7010926）

介绍：

Benefiting from facile data acquisition, DECAF comprises affective responses of 30 subjects to 36 movie clips and 40 one-minute music video segments (used in [15]), making it one of the largest available emotional databases.

[15] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebra- € himi, T. Pun, A. Nijholt, and I. Patras, “DEAP: A database for emotion analysis using physiological signals,” IEEE Trans. Affective Comput., vol. 3, no. 1, pp. 18–31, Jan.-Mar. 2012.

视频剪辑：

This section describes how the 36 movie clips compiled to this end were selected. Based on previous studies that have identified movie clips suited to evoke various target emotions [2], [9], we initially compiled 58 Hollywood movie segments. These clips were shown to 42 volunteers, who self-assessed their emotional state on viewing each video to provide: valence level (very negative to very positive), arousal level (very calm to very excited), and the most appropriate tag that describes the elicited emotion.

The chosen movie clips were 51.1-128.2 s long (m = 80; s = 20) and were associated with diverse emotional tags. For benchmarking affective stimuli, we also recorded emotional responses to 40 one-minute music video used in the DEAP study [15].

[2] E. E. Bartolini, “Eliciting emotion with film: Development of a stimulus set,” Master’s thesis, Wesleyan University, Connecticut, USA, 2001.

[9] J. J. Gross and R. W. Levenson, “Emotion elicitation using films,” Cognition Emotion, vol. 9, no. 1, pp. 87–108, 1995.

