Computational methods for a customised positive mood-supporting system based on multi-sensorial stimuli

Claudia Rabaioli[1] (PhD student)

[1]Dept. of Informatics, Systems and Communication DISCo, University of Milano-Bicocca, Milano, Italy

Abstract

The following project concerns the proposal of a positive mood-supporting application, based on multi-sensorial stimuli. The first goal of this project is to study the differences, relations, and interactions between emotion and mood, trying to understand how emotion recognition methods and eliciting stimuli could be adjusted in the mood domain. This is a fundamental step in building a system meant to detect the user's mood and monitor it over time while trying to support a positive mood by presenting proper multisensorial stimuli. Therefore, two other intermediate steps are essential: i) building a multimodal mood detection framework, using wearable devices combined with short questionnaires; and ii) defining which audiovisual characteristics a multi-sensorial stimulus should exhibit to empower its positive mood-support effect, exploiting different learning models.

A profiling questionnaire will be submitted the first time a user accesses the system. Then a multimodal mood detection takes place with an Ecological Momentary Assessment (EMA) of the current mood and data derived from smartphones and wearable devices such as environmental, behavioural, and physiological data. Since the user's current mood is defined, the system automatically selects the proper audiovisual stimuli to suggest the improvement of the mood, if needed. After a mood support session, a user's feedback will be required to improve the effectiveness of the system and ensure a better set of stimuli for the specific user. A case study will be considered to understand the efficacy of this system in real-world applications. In particular, an automotive scenario will be developed, exploiting virtual reality stimuli during a driving simulation.

Keywords

mood detection, multi-sensorial stimuli, multimodal approach, hand-crafted features

1. Introduction

Emotion recognition is a widely investigated field [1]: emotion is a programmed but instantaneous neural response. Mood is, instead, the expression of an affective state over time. Current mood can influence daily life, for example, general well-being, work productivity, and how a human being reacts in stressful situations. Despite its importance, mood has no common definition in the literature and is often mistaken or investigated as an emotion, feeling, or affective state. Its dominant characteristic is its persistence over time and its unrelatedness to a specific stimulus, unlike emotions and feelings [2, 3, 4, 5].

In the psychiatric disorders field [6, 7], recent studies introduce Ecological Momentary Assessment (EMA) as an alternative to static retrospective reports [8], allowing subjects to report real-time experiences in a real-world scenario, repeatedly over time, apprehending the mood flow due to the context and events [9].

Smartphones can meet the momentary necessity of this type of assessment, due to their handy nature [10]; they can also provide, together with other wearable and portable devices, physiological, behavioural, environmental, and contextual information. Following the suggestions presented by Pace-Schott et al. [11], the study of the difference between mood and emotion can start from the use of EMA questionnaires and physiological data. This hypothesis may define a mood detection protocol by adjusting the emotion recognition systems. Recently, patients affected by mood disorders have also been studied with the support of wearable and portable devices that allow passive sensing [12], collecting data without user effort [13] or conditioning in real-life scenarios. Smartphones can dispose of different sensors, such

Doctoral Consortium at the 23rd International Conference of the Italian Association for Artificial Intelligence Bolzano, Italy, November 25-28, 2024

claudia.rabaioli@unimib.it (C. Rabaioli)

as microphones, accelerometers, and GPS, useful to perform speech emotion recognition, physical activity monitoring, and the user's location assessment, to access meteorological data, and air quality, and to distinguish between work, travelling, and rest environments. Smartwatch sensors can collect physiological data like cardiac-related ones, useful for emotion detection, mood variations, sleep, and physical activity monitoring.

The available mood self-reports are mainly based on or taken directly from the questionnaires designed for related fields such as emotion, affect, and personality [14]. Verbal questionnaires, such as Differential Emotions Scale (DES) [15], Profile of Mood States (POMS), and Positive and Negative Affect Schedule (PANAS) [16], ask participants to rate emotional items (e.g., adjectives) on a 5-point scale, expressing the similarity to the current state or the frequency of the experience in the last period. The items of FACES test [17] and Pick-A-Mood [14] are instead visual representations of different moods, demanding users to pick the stylized face that is more similar to their current mood.

Audiovisual stimuli can help to positively support mood, playing a part in determining a quieter state of a subject. The available datasets are intended to elicit emotions, or affect a user state, and are not specifically designed to influence moods [18]. They include unimodal stimuli, for example, International Affective Picture System (IAPS) [19], Geneva Affective PicturE Database (GAPED) [20], Nencki Affective Picture System (NAPS) [21], EmoPics [22], International Affective Digitized Sounds (IADS) [23], except for Experimental MOVies for Induction of Emotions (E-MOVIE) [24], which employs audio and visual stimuli.

2. Proposal

My project proposal regards a system intended to detect a user's current mood and automatically offer multi-sensorial stimuli to support a positive influence on the user's state. To detect the mood, considering its definition, daily, long-term, and non-invasive monitoring of the user's state seems essential. An optimal solution would be a smartphone application, that also employs wearable devices, in order to collect information about the user comprising self-assessment reports and the analysis of recorded multimodal data.

A machine learning model feed with hand-crafted features will be developed to study which kind of multi-sensorial stimuli offer the best positive support and which stimuli features should be included in audiovisual production best practices. This system will leverage a personalised approach by implementing user-adaptive models for mood detection and stimuli suggestion.

The proposed system can be employed in different fields with the aim of improving user's well-being. For example, a more favourable ambiance built with positive mood support multi-sensorial stimuli can support the psychological well-being of patients during rehabilitation or senior citizens. Also, the automotive field can benefit from the proposed system, where driving simulations can measure the support effect in stressful situations.

Interpreting the proposed system as a potential medical assistance item, the European AI Act regulations will be followed thoroughly, especially considering the system potentiality of reducing stress in working environments. This application would also be in line with the 2030 Agenda for Sustainable Development of the United Nations and in particular with the goal of improving healt and well-being in work environments. Ethical considerations are clearly paramount. For example, it will be essential to ensure user anonymity during data processing, with pre-processed data stored on a cloud server. A rigorous validation process will be implemented to ensure a positive impact on users' moods. Additionally, user volunteers will be observed, allowing them to determine their benefits from the support system. Legal experts will be consulted to ensure compliance with the EU Artificial Intelligence Act and the General Data Protection Regulation (GDPR).

To realize the final system design, several steps must be addressed: each of them is categorized into distinct topic blocks, as shown in Figure 1. Each block represents an innovative approach and will be accompanied by specific milestones to track progress.

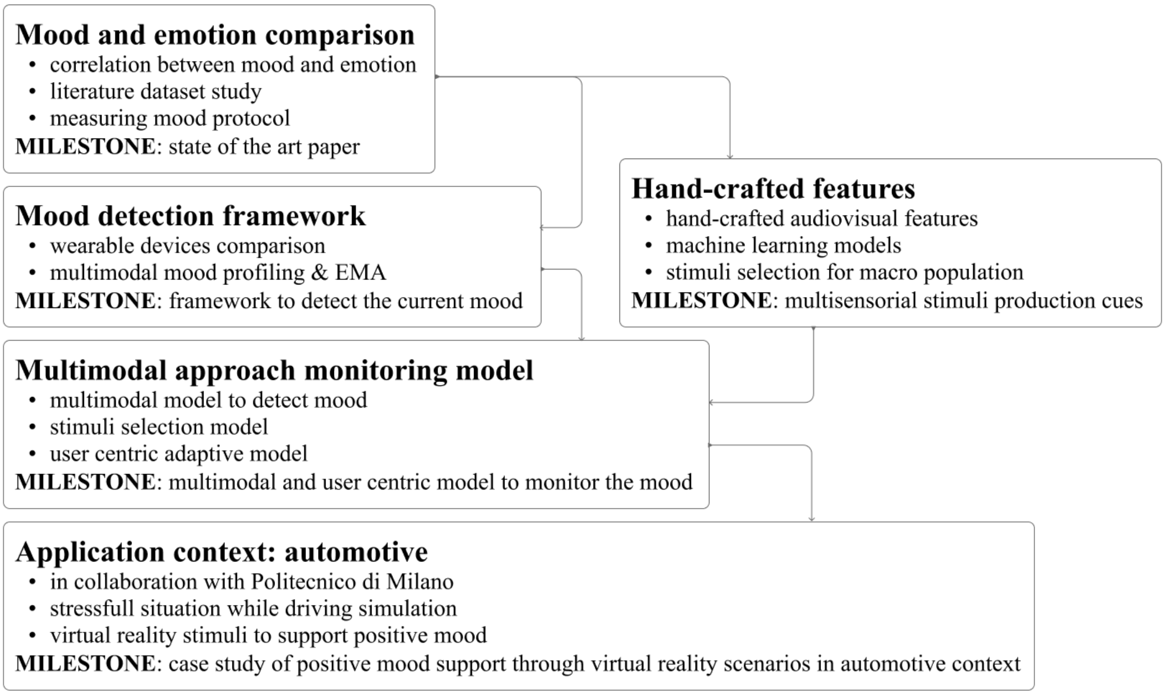

Figure 1: Diagram of topics and milestones of the project proposal

2.1. Mood and Emotion comparison

Firstly, a deep study of the state of the art will be carried out to better understand mood and emotion. Their influence and dependence on each other are open questions and need to be further investigated. The main aim is to understand if (and how) we can study mood leveraging on emotional response to stimuli and if it is possible to employ emotion recognition models during long periods of user observation. Moreover, it is important to understand if the emotionally eliciting stimuli can be used to support a positive mood, as long as it is repeated over time. This will require the study of different emotion stimuli datasets and the selection of suitable starting stimuli. After this initial step, a protocol to measure the mood will be defined and the preparation of a state-of-the-art paper is intended to conclude this first phase.

2.2. Mood detection framework

A smartphone is hypothesized to be a suitable tool to support the definition of a framework to detect a user's current mood, allowing the development of an ecological system for passive monitoring. This handy and real-life approach would be a new way to detect mood for common users, considering that previous studies were mainly conducted on clinical cases, such as patients with mood disorders, in controlled contexts.
As previously introduced, the system includes multimodal mood profiling, thus the application will require the following data at registration: demographics, such as age, gender, education level, and the predominant personality trait, by answering the 10-item Big Five Inventory Personality Test [25]. Subsequently, the current mood will be accessed through an EMA questionnaire, exploiting the Pick-A-Mood validated questionnaire.
Moreover, environmental and contextual (e.g., weather and location), behavioural (e.g., facial expression and speech analysis [26]), and physiological data (e.g., photoplethysmographic signals and physical activity [27]) will be collected through wearable and portable devices such as smartphones and smartwatches. The type of collected data strongly depends on the type of available wearable device, thus a comparative analysis between different devices will be performed.

2.3. Multisensorial stimuli characterization

The literature will be exploited to find suitable handcrafted features and learning models to use on the previously identified literature datasets and to determine which audiovisual characteristics are mostly responsible for the positive effect of the multi-sensorial stimuli. With positive and negative effect labelling, a supervised learning model will be applied to predict the classes efficiently and determine the best-ranked features during the classification.

The model will use carefully selected handcrafted features based on humanistic approaches instead of relying on less interpretable deep features. In cinematography, best practices help create engaging content, which can be translated into interpretable handcrafted features. This study builds on my previous thesis, where I explored integrating humanistic and psychological perspectives into computer vision algorithms. By analyzing these features, we can create production guidelines for developing updated datasets across various fields, focusing on areas like communication effectiveness and interestingness, as well as supporting positive mood, which is the aim of this project.

For this project, specific stimuli will be chosen to define a general dataset for a population, to improve the effectiveness of the positive mood support. This study in the future will allow the creation of new datasets, with updated styles or specially made for specific types of populations. The acquired cues will aid in creating datasets with new types of stimuli, such as tridimensional environments for virtual reality, necessitating user participation for data validation.

2.4. Multimodal approach monitoring

In affective computing literature, multimodal models are widely used to recognize emotions [1]. Combining behavioural data, such as speech, facial expressions, and body movements, with physiological parameters has been demonstrated to improve the emotion recognition capability of different models [28]. The aim of this phase is the creation of a mood detection system, starting from emotion recognition models to be translated into a continuous mood monitoring of the system users.

The feature set and learning model will be carefully selected in previous steps to ensure that the application can provide real-time responses. Common problems from systems that integrate heterogeneous data from different sources will be considered. This structure will also adapt to individual users to support a user-centric approach, not only in detecting current moods but also in suggesting appropriate stimuli. The model will offer positive support stimuli to enhance emotional well-being. The effectiveness of these stimuli will be measured by monitoring mood changes over time and gathering feedback from users.

2.5. Possible application scenario

I am currently collaborating with the Haptics and Virtual Prototyping Lab in the Department of Mechanical Engineering at Politecnico di Milano, under the direction of Professor Francesco Ferrise. This research group primarily focuses on virtual and augmented reality, with active involvement in the automotive field. We will test the mood support system by creating virtual reality scenarios that simulate stressful driving situations. In these scenarios, specific audiovisual stimuli could serve as effective tools to reduce driver tension and promote safer driving conditions. Research has shown that daily stress is associated with negative mood changes on the same day [29, 30].

Through this application scenario, we will evaluate the efficacy of the system, considering various metrics to identify the most appropriate ones for the specific final application.

3. Research contribution

Computational models play an important role in this project: an innovative mood profiling and detection system would integrate multimodal data, such as personality portray, recently studied EMA questionnaires, and environmental, contextual, behavioural, and physiological data thanks to portable

and wearable devices.
Using consumer wearable devices has the advantage of being economically available to extended groups of people, but also the weakness of less precise data [31]; complementing those data with EMA and different kinds of data, not only the psychological ones, will help with the accuracy of mood detection. Multimodal approaches are adopted in the emotion recognition field; this study would also allow us to understand the correlations between mood and emotions over time. The handy approach permits the investigation of subjects with no particular pathologies while living in real-world situations and not in controlled environments.
Machine learning models will also help to characterize the most engaging stimuli, thanks to hand-crafted features designed from the video-making best practices. The obtained creative cues will also help to create new datasets in the future and improve virtual scenarios such as in driving simulation.
Computational models will also offer challenges to overcome, such as a user-centric approach for adaptive mood detection, suitable efficiency for a smartphone application, and collecting data from different devices.

References

[1] Y. Cimtay, E. Ekmekcioglu, S. Caglar-Ozhan, Cross-subject multimodal emotion recognition based on hybrid fusion, IEEE Access 8 (2020) 168865–168878. doi:10.1109/ACCESS.2020.3023871.

[2] M. Siemer, Moods as multiple-object directed and as objectless affective states: An examination of the dispositional theory of moods, Cognition & Emotion 19 (2005) 815–845.

[3] C. N. Moridis, A. A. Economides, Mood recognition during online self-assessment tests, IEEE Transactions on Learning technologies 2 (2009) 50–61.

[4] C. J. Beedie, P. C. Terry, A. M. Lane, T. J. Devonport, Differential assessment of emotions and moods: Development and validation of the emotion and mood components of anxiety questionnaire, Personality and Individual Differences 50 (2011) 228–233.

[5] S. Saganowski, B. Perz, A. Polak, P. Kazienko, Emotion recognition for everyday life using physiological signals from wearables: A systematic literature review, IEEE Transactions on Affective Computing (2022).

[6] L. Hollands, J. Lambert, L. Price, D. Powell, C. Greaves, Ecological momentary assessment of mood and physical activity in people with depression, Journal of affective disorders 271 (2020) 293–299.

[7] M. Hall, P. V. Scherner, Y. Kreidel, J. A. Rubel, A systematic review of momentary assessment designs for mood and anxiety symptoms, Frontiers in Psychology 12 (2021) 642044.

[8] A. A. Stone, S. Shiffman, Ecological Momentary Assessment (Ema) in Behavioral Medicine, Annals of Behavioral Medicine 16 (1994) 199–202. URL: https://doi.org/10.1093/abm/16.3.199. doi:10.1093/abm/16.3.199.

[9] S. Shiffman, A. A. Stone, M. R. Hufford, Ecological momentary assessment, Annu. Rev. Clin. Psychol. 4 (2008) 1–32.

[10] H. Torkamaan, J. Ziegler, Exploring chatbot user interfaces for mood measurement: a study of validity and user experience, in: Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers, 2020, pp. 135–138.

[11] E. F. Pace-Schott, M. C. Amole, T. Aue, et al, Physiological feelings, Neuroscience & Biobehavioral Reviews 103 (2019) 267–304. doi:10.1016/j.neubiorev.2019.05.002.

[12] M. B. Morshed, K. Saha, R. Li, S. K. D'Mello, M. De Choudhury, G. D. Abowd, T. Plötz, Prediction of mood instability with passive sensing, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 3 (2019) 1–21.

[13] V. P. Cornet, R. J. Holden, Systematic review of smartphone-based passive sensing for health and wellbeing, Journal of biomedical informatics 77 (2018) 120–132.

[14] P. M. A. Desmet, M. H. Vastenburg, N. Romero, Mood measurement with pick-a-mood: Review of

current methods and design of a pictorial self-report scale, Journal of Design Research 14 (2016) 241–279. doi:10.1504/JDR.2016.079751.

[15] C. E. Izard, P. B. Read, Measuring emotions in infants and children: based on seminars sponsored by the Committee on Social and Affective Development During Childhood of the Social Science Research Council, volume 1, Cambridge University Press, 1982.

[16] D. Watson, L. A. Clark, A. Tellegen, Development and validation of brief measures of positive and negative affect: the panas scales., Journal of personality and social psychology 54 (1988) 1063.

[17] G. Puertas, V. Patel, T. Marshall, Are visual measures of mood superior to questionnaire measures in non–western settings?, Social psychiatry and psychiatric epidemiology 39 (2004) 662–666.

[18] A. Alnafjan, M. Hosny, Y. Al-Ohali, A. Al-Wabil, Review and classification of emotion recognition based on eeg brain-computer interface system research: A systematic review, Applied Sciences 7 (2017) 1239. doi:10.3390/app7121239.

[19] P. J. Lang, M. M. Bradley, B. N. Cuthbert, et al., International affective picture system (iaps): Technical manual and affective ratings, NIMH Center for the Study of Emotion and Attention 1 (1997) 3.

[20] E. S. Dan-Glauser, K. R. Scherer, The geneva affective picture database (gaped): a new 730-picture database focusing on valence and normative significance, Behavior research methods 43 (2011) 468–477.

[21] A. Marchewka, Ł. Żurawski, K. Jednoróg, A. Grabowska, The nencki affective picture system (naps): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database, Behavior research methods 46 (2014) 596–610.

[22] M. Wessa, P. Kanske, P. Neumeister, K. Bode, J. Heissler, S. Schönfelder, et al., Emopics: Subjektive und psychophysiologische evaluation neuen bildmaterials für die klinisch-bio-psychologische forschung, Zeitschrift für Klinische Psychologie und Psychotherapie 39 (2010) 77.

[23] M. M. Bradley, P. J. Lang, International affective digitized sounds (iads): Stimuli, instruction manual and affective ratings (tech. rep. no. b-2), Gainesville, FL: The Center for Research in Psychophysiology, University of Florida (1999).

[24] A. Maffei, A. Angrilli, E-movie-experimental movies for induction of emotions in neuroscience: An innovative film database with normative data and sex differences, Plos one 14 (2019) e0223124.

[25] R. R. McCrae, O. P. John, An introduction to the five-factor model and its applications, Journal of personality 60 (1992) 175–215.

[26] R. Jahangir, Y. W. Teh, F. Hanif, G. Mujtaba, Deep learning approaches for speech emotion recognition: state of the art and research challenges, Multimedia Tools and Applications 80 (2021) 23745–23812.

[27] S. Wan, L. Qi, X. Xu, C. Tong, Z. Gu, Deep learning models for real-time human activity recognition with smartphones, Mobile Networks and Applications 25 (2020) 743–755.

[28] K. Ezzameli, H. Mahersia, Emotion recognition from unimodal to multimodal analysis: A review, Information Fusion 9 (2023). doi:10.1016/j.inffus.2023.101847.

[29] A. DeLongis, S. Folkman, R. S. Lazarus, The impact of daily stress on health and mood: psychological and social resources as mediators., Journal of personality and social psychology 54 3 (1988) 486–95.

[30] A. A. Stone, J. M. Neale, S. Shiftman, Daily assessments of stress and coping and their association with mood, Annals of Behavioral Medicine 15 (1993) 8–16. URL: https://doi.org/10.1093/abm/15.1.8. doi:10.1093/abm/15.1.8.

[31] S. Gashi, P. Oldrati, M. Moebus, M. Hilty, L. Barrios, F. Ozdemir, P. Consortium, V. Kana, A. Lutterotti, G. Rätsch, et al., Modeling multiple sclerosis using mobile and wearable sensor data, npj Digital Medicine 7 (2024) 64.