

Prerequisites for Affective Signal Processing (ASP) – Part IV

Egon L. van den Broek¹, Marjolein D. van der Zwaag², Jennifer A. Healey³
Joris H. Janssen^{2,4} and Joyce H. D. M. Westerink²

¹ Human-Centered Computing Consultancy, Austria

² User Experience Group, Philips Research Europe
High Tech Campus 34, 5656 AE Eindhoven, The Netherlands

³ Future Technology Research, Intel Labs Santa Clara
Juliette Lane SC12-319, Santa Clara CA 95054, U.S.A.

⁴ Department of Human Technology Interaction, Eindhoven University of Technology
P.O. Box 513, 6500 MB Eindhoven, The Netherlands

Abstract. In [1–3], a series of prerequisites for affective signal processing (ASP) was defined: validation (e.g., mapping of constructs on signals), triangulation, a physiology-driven approach, contributions of the signal processing community, identification of users, theoretical specification, integration of biosignals, and physical characteristics. This paper defines three additional prerequisites: historical perspective, temporal construction, and real-world baselines.

1 Introduction

In his book *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind*, [4] stated: ... *emotion is one of those suitcaselike words that we use to conceal the complexity of very large ranges of different things whose relations we don't yet comprehend*. Five pages later, he suggests to replace ... *old questions like, "What sorts of things are emotions and thoughts?" by more constructive ones like, "What processes does each emotion involve?" and "How could machines perform such processes?"* Affective computing (AC) aims to answer these questions through processing signals that correlate with emotions: affective signals processing (ASP).

ASP can be employed from (a combination of) biosignals, movement analysis, computer vision, and speech processing. However, the techniques other than biosignals have major disadvantages [1–3]. In contrast, such issues have been resolved for biosignals in recent years: currently, it is easy to obtain, high fidelity, cheap, and unobtrusive biosignal recordings; e.g., see [5]. Moreover, the recording devices can be easily integrated in various products [6]. Therefore, this paper focusses on biosignals. For an overview of the most commonly used biosignals and their features, we refer to [1].

This prerequisites paper is designed to discuss unsolved issues related to ASP and to introduce a framework for future research. It is not designed as a paper on novel methods in signal processing, but rather on the specific issues on applying those methods to the problem of ASP. A particular focus is on the problem of ASP in the real world, with long latency signals (e.g., electrodermal activity; EDA), and affective responses that are ambiguously defined in time and that often depend on previous events and are, therefore, neither linear nor time invariant in their responses. Much of (traditional) signal processing relies on the linear time invariant assumption. Real affective responses do not fit this description. Consequently, ASP requires its own set of prerequisites as they are denoted in this paper and the other prerequisites papers of [1–3].

For AC, a broad plethora of classifiers is used as part of the ASP. The classification performances are hard to compare since the emotion classes used are typically defined in different ways. Additionally, the number of emotion classes to be discriminated is small, it ranges from 2 to 6. Nevertheless, the results are behind that of other classification problems. With AC recognition rates $\ll 90\%$ are common, where in most other pattern recognition problems, recognition rates of $> 90\%$ and often $> 95\%$ are often reported. This illustrates AC's complex nature and the need for a comprehensive review of the prerequisites involved.

To force a breakthrough in results on AC we propose a set of prerequisites for ASP, before starting with AC in practice. The first three parts of these prerequisites were introduced in [1–3]. In the next section, the fourth part is introduced. Together, these prerequisites should form the foundation for more successful ASP and AC. We end this paper with a brief conclusion.

2 Prerequisites – Part IV

In [1–3], the following prerequisites for ASP were introduced: validity, triangulation, a physiology-driven approach, contributions from signal processing, user identification, theoretical specification, integration of biosignals, and physical characteristics. While each of these is still of the utmost importance for ASP, we will now denote three additional ones: historical perspective, temporal construction, and real-world baselines.

2.1 History: Lessons to be Learned and Experiences to Remember

Centuries ago, the relation between physiological reactions, as expressed through biosignals, and emotions was already mentioned by poets and ancient philosophers. This resulted in a plethora of definitions, almost impossible to list and illustrates the complexity of the concept emotion; cf. [4].

Although much knowledge on emotions is gained over the last centuries, researchers tend to ignore this up to a high extent and stick to some relatively recent theories; e.g., the valence and arousal model or the approach avoidance model. This holds in particular for ASP and AC, where an engineering approach is dominant and a theoretical framework is considered of lesser importance [2]. Consequently, for most engineering approaches, the valence-arousal model is applied as a default option, without considering other possibilities.

It is far beyond the scope of this paper to provide a complete overview of all literature relevant for ASP and AC. For such an overview, we refer to the various handbooks and review papers on emotions, affective sciences, and affective neuroscience; e.g., [7–9]. In this section, we will touch some of the major works on emotion research, which origin from medicine, biology, physiology, and psychology.

Let us start with one of the earliest works on biosignals: *De l'Électricité du corps humain* of M. l'Abbé Bertholon (1780), who already described human biosignals. One century later Darwin (1872) published his book *The Expression of Emotions in Man and Animals* [8]. Subsequently, independently of each other, William James and C. G. Lange revealed their theories on emotions, which were remarkably similar [8]. Consequently, their theories has been merged and were baptized the James-Lange theory.

In a nutshell, the James-Lange theory argues that the perception of our own biosignals *is* the emotion. Consequently, no emotions can be experienced without these biosignals. Two decades after the publication of James' theory, it was already seriously challenged by [11, 12] and [13, 14]. They emphasized the role of subcortical structures (e.g., the thalamus, the hypothalamus, and the amygdala) in experiencing emotions. Their rebuttal on the James-Lange theory was founded on five notions:

1. Compared to a normal situation, experienced emotions are similar when biosignals are omitted; e.g., as with the transection of the spinal cord and vagus nerve.
2. Similar biosignals emerge with all emotions. So, these signals cannot cause distinct emotions.
3. The bodies internal organs have fewer sensory nerves than other structures. Hence, people are unaware of their possible biosignals.
4. Generally, biosignals have a long latency period, compared to the time emotional responses are expressed.
5. Drugs that trigger biosignals to emerge do not necessarily trigger emotions in parallel.

We will now address each of Cannon's notions from the perspective of ASP. As will become apparent, considering these notions with current ASP is of importance. To the authors knowledge, the first case that illustrated both theories weaknesses was that of a patient with a lesion, as denoted in Cannon's first notion. This patient reported : *Sometimes I act angry when I see some injustice. I yell and cuss and raise heel, because if you don't do it sometimes, I learned people will take advantage of you, but it just doesn't have the heat to it that it used to. It's a mental kind of anger* (p. 151) [15]. Moreover, this case clearly illustrated the use of such special cases, as is denoted in [2].

The second notion of the Cannon-Bard theory strikes the essence of ASP. It would imply that the quest of affective computing is deemed to fail. According to Cannon-Bard, ASP is of no use since no unique sets of biosignals exist that map to distinct emotions. Luckily, nowadays, this statement is judged as coarse [8]. However, it is generally acknowledged that it is very hard to apply ASP successfully [7]. So, (at least) to a large extent Cannon was right.

It was confirmed that the number of sensory nerves differs in distinct structures in human bodies (Cannon's notion 3). So, indeed people's physiological structures determine their internal variations to the emotional sensitivity. To make ASP even more

challenging, cross-cultural and ethnic differences exist in people's patterns of biosignals, as was already shown by [16]⁵.

The fourth notion concerns the latency period of biosignals, which Cannon denoted as being 'long'. In the next section we address this problem.

The fifth and last notion of Cannon is one that is not addressed so far. It goes beyond biosignals since it concerns the neurochemical aspects of emotions. Although this component of human physiology can indeed have a significant influence on experienced emotions, this falls far beyond the scope of this paper.

It should be noted that the current general opinion among neuroscientists is that the truth lies somewhere between the theories of James-Lange and Cannon-Bard [8]. However, the various relations between the latter notions and the set of prerequisites, illustrates that these notions, although a century old, are still of interest for current AC and ASP.

2.2 Temporal Construction

There are many temporal aspects in biosignals that should be taken into account in ASP. These aspects can be categorized in three classes: psychological, physiological, and signal processing aspects.

The psychological aspect has to do with habituation; in general, every time a stimulus is perceived one's reaction to it will get smaller. With large delays between the stimuli, one recovers from the habituation effect. There are several ways of dealing with this in ASP. One way is to keep track of the moments in which stimuli were present. This information can then be used to predict how strong the effect of a similar stimulus will be. Alternatively, in applications where stimuli presentation can be controlled, the variety of the stimuli can be directed such that habituation effects are canceled.

The first physiological aspect deals with the fact that the affective signals can be processed in different time windows. For instance, we can look at parts of 30 minutes but also at 10, 30, or 60 seconds. There are many challenges in modeling the temporal aspects of emotion. One is the annotation challenge of determining when the emotion begins and when it ends. Another is the sensor fusion problem of determining how to window individual signals within the emotional event since different signals have different latencies. In response to a high arousal event, an instant gasp may occur in respiration and a tensing of muscles, heart rate will then increase in the next few seconds and EDA should start to rise and may continue to rise for several minutes. Using the same window and offset for all signals would not capture the most salient discriminating features of the experience. So, in general, biosignal features calculated over time windows with different length cannot be compared with each other.

That being said, time window selection is often done empirically; i.e., many different time windows are tried and those leading to the best results are used in the final models [17, 6]. Other automatic options include finding the nearest significant local minima or making assumptions about the start time and extend of the emotion; e.g., an average over previous emotions. In addition, another empirical solution is to ask the

⁵ *Author's note.* Nowadays, this paper would run up to resistance, as it denotes both ethnical issues and as its subjects. Perhaps that is why so little work is done on this topic.

user to define the window of interest. This can be done through sliders, as for real world research can be presented on a PDA. However, also this approach has its downside: in general, people's introspection is not good and you do not want to bother users with these tasks. The physiological response to an emotion may have started well before the person realized that they were in this state, so if a single annotation is used, it will definitely come after the start of the experience. Moreover, in the real world the temporal nature of the reaction to the stimulus is undetermined. A uniform window may not be appropriate.

There are also valuable theoretical considerations. Different psychological processes develop over different time scales. On the one hand, emotions lead to very short and fast phasic changes and, thus, require short time windows. On the other hand, changes in mood are more gradual tonic and, so, require broader time windows. In general, the time window used should depend on the psychological construct studied. Furthermore, there is always a lag between the psychological change and the physiological change. These lags differ per signal: heart rate changes almost immediately while skin temperature can take more than a minute to change. Skin conductance is somewhere in between. This shows the need for different time windows for different signals.

A second physiological aspect stems from the idea that physiological activity tends to move to a stable neutral state; i.e., when the physiological level is high, it tends to decrease; whereas, when the physiological level is low, it tends to increase. Hence, the effect of a stimulus on physiology depends on the physiological level before stimulus onset; i.e., the principle of initial values [9]. When you perceive a scary stimulus and your heart rate is at 80 it might increase by 15 beats, however, when your heart rate is at 160 it is unlikely to increase at all. As this is found to be a linear relationship, it can be modeled by linear regression. The first step is to assess the regression line, which is different per feature and person. Next, this regression line can be used to correct each feature by computing its residualized value.

A consideration specific to ASP is that emotional responses are likely comprise a layered response involving components that have different time periods including: disposition (long term - years), circumstance (days), mood (hours) and emotion (seconds). An accurate model of an individual's affective response to these varying time influences is difficult to determine, even the totality of influences are difficult to catalog in the real world. Therefore, also for this reason, a major consideration in ASP is choosing a window length appropriate to the type of affective response you are considering.

2.3 Real-world Baselines

Baselining is the process of correcting the biosignal to a standard level that is comparable over users and/or sessions (also called normalization/standardization). Finding an appropriate baseline is both important and difficult for sensors whose readings depend on factors that can easily change on a daily basis; e.g., sensor placement, humidity, temperature, and the use of contact gel [3]. Still, baselining over multiple people or multiple days is required for ASP in order to compare and combine data from these different sources in a meaningful way. Many different approaches to baselining are known in the literature. However, as will be shown, we require affect specific approaches.

Table 1. Seven methods of using baseline information to normalize the signal. x denotes the original signal and \tilde{x} the corrected signal. μ_B , \min_B , \max_B , and σ_B are respectively the mean, minimum, maximum, and standard deviation of the baseline. Sources of information: 1:[18], 3,5,6:[19] and 7:[20].

1 $\tilde{x}_i = x_i - \mu_B$	Standard correction, often used in psychological experiments.
2 $\tilde{x}_i = x_i - \min_B$	Useful alternative to the first method when there is no relaxation period and a lot of variance in the signal.
3 $\tilde{x}_i = (x_i - \mu_B) / \sigma_B$	Strong baselining method; works best for continuous signals.
4 $\tilde{x}_i = (x_i - \mu_B) / \mu_B$	
5 $\tilde{x}_i = (x_i - \min_B) / (\max_B - \min_B)$	Sensitive to outliers.
6 $\tilde{x}_i = x_i / \max_B$	Used for Skin conductance responses features.
7 $\tilde{x}_i = (x_i \times 100) / \mu_B - 100$	Used for facial EMG measurements.

With ASP we have to handle long term continuous (re-)baselining of biosignals. There exists no guideline on how to apply these methods to continuous physiological data in the real world. An exception to this is [21], which discusses ECG recording in ambulatory settings; however, it does not focus on ASP. In this section, we discuss how to apply known methods from the laboratory to a new situation, continuous ambulatory monitoring in the real world. We also discuss how to apply these methods to affective reactions of varying length and intensity in the presence of noise. Some of the methods commonly used in other types of signal processing, such as “zeroing the mean” and “dividing by the variance” do not work for long term physiological records, which is the problem we are trying to bring to light. In the following paragraphs, we will try to give an overview of the different baselining approaches and explain when they are appropriate. We also call for empirical comparisons of different baseline methodologies specific to ASP in the real world, as this is still lacking.

The two main issues with baselining are (1) the selection of a suitable correction method and (2) the selection of a period over which to calculate the parameters of the selection method (the baseline period). The correction methods are summarized in Table 1. Once the baseline is removed, it becomes the new base (or zero) and the original value is lost. Each baseline has different merits. Taking the minimum baseline is more equivalent to taking the resting EDA that would normally be used in a laboratory experiment. This is the best method if a consistent minimum seems apparent in all data being combined. The problem is that for each data segment, a minimum must be apparent. It is straightforward to eliminate point outliers such as those at 3.7 hours and 3.9 hours and find a more robust minimum for the baseline. An other often used method for continuous signals like EDA and skin temperature is called standardization (method 3 in Table 1) [19]. This is probably the most powerful correction method and is applied very often. It corrects not only for the baseline level but also for the variation in the signal, making it more robust. Other correction methods are tailored to specific features; e.g., the amplitude of skin conductance responses is often corrected by dividing by the maximum amplitude. Taken together, different signals and situations require different correction methods, which should be chosen carefully.

The second issue in baselining is the selection of an appropriate time-window over which the parameters for the correction are calculated. For short term experiments, a single baseline period is usually sufficient. However, when monitoring continuously,

the baseline may have to be re-evaluated with greater frequency. The challenge here is to find a good strategy for dividing the signal into segments over which the baseline should be re-calculated. A simple solution is to use a sliding window; e.g., where the last 30 minutes are taken into account. In this case, it seems obvious that the segments should be considered independently, since the time period between the two is long and it may be that the electrodes fell off and may have been re-applied. To be able to conduct proper normalization when signal loss is shorter than the baselining window, the period right before the signal loss can for example be used to complement the baseline window of the current signal. However, in general, data segmentation has not had much attention in ASP and long term continuous (re-)baselining of biosignals is still an open problem.

As with all pattern recognition pipelines, baselining (or normalization) is of utmost importance. Most efforts towards AC and ASP have not paid much attention to this. We hope that this prerequisite is a start for the development of more sophisticated algorithms that can deal with the difficult problems of data segmentation, as they have been dealt with in other research fields like computer vision.

3 Conclusions

This paper explains the importance of prerequisites specifically for ASP and introduces the fourth part of a series of such prerequisites for ASP. The prerequisites foundation in historical perspective, adequate temporal construction, and well-chosen real-world baselines are introduced. These prerequisites are complementary to those presented in [1–3]: validity, triangulation, the physiology-driven approach, and contributions of signal processing, identification of users and theoretical specification, and physical characteristics and integration of biosignals.

The review and the prerequisites, both illustrate and explain the complexity of ASP and its limited progress. Therefore, we advise to incorporate these prerequisites for successful ASP, instead of running forward and ignoring the problems encountered in previous studies. We hope that the prerequisites can contribute to or even guide future research in ASP.

References

1. van den Broek, E.L., Janssen, J.H., Westerink, J.H.D.M., Healey, J.A.: Prerequisites for Affective Signal Processing (ASP). In Encarnação, P., Veloso, A., eds.: Biosignals 2009: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, Porto – Portugal (2009) 426–433
2. van den Broek, E.L., Janssen, J.H., Healey, J.A., van der Zwaag, M.D.: Prerequisites for Affective Signal Processing (ASP) – Part II. In: Biosignals 2010: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, Valencia – Spain (2010) [in press]
3. van den Broek, E.L., Janssen, J.H., van der Zwaag, M.D., Healey, J.A.: Prerequisites for Affective Signal Processing (ASP) – Part III. In: Biosignals 2010: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, Valencia – Spain (2010) [in press]

4. Minsky, M.: *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York, NY, USA: Simon & Schuster (2006)
5. Pantelopoulos, A., Bourbakis, N.G.: A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 40 (2010) 1–12
6. van den Broek, E.L., Westerink, J.H.D.M.: Considerations for emotion-aware consumer products. *Applied Ergonomics* 40 (2009) 1055–1064
7. Boehner, K., DePaula, R., Dourish, P., Sengers, P.: How emotion is made and measured. *International Journal of Human-Computer Studies* 65 (2007) 275–291
8. Dalglish, T., Dunn, B.D., Mobbs, D.: Affective neuroscience: Past, present, and future. *Emotion Review* 1 (2009) 355–368 history.
9. Davidson, R.J., Scherer, K.R., Hill Goldsmith, H.: *Handbook of affective sciences*. New York, NY, USA: Oxford University Press (2003)
10. l'Abbé Bertholon, M.: *De l'Électricité du corps humain*. Lyon, France: Tome Premiere (1780)
11. Cannon, W.B.: *Bodily changes in pain, hunger, fear and rage: An account of recent researches into the function of emotional excitement*. New York, NY, USA: D. Appleton and Company (1915)
12. Cannon, W.B.: The James-Lange theory of emotion: A critical examination and an alternative theory. *American Journal of Psychology* 39 (1927) 106–124
13. Bard, P.: On emotional expression after decortication with some remarks on certain theoretical views, Part I. *Psychological Review* 41 (1934) 309–329
14. Bard, P.: On emotional expression after decortication with some remarks on certain theoretical views, Part II. *Psychological Review* 41 (1934) 424–449
15. Hohnmann, G.W.: Some effects of spinal cord lesions on experienced emotional feelings. *Psychophysiology* 3 (1966) 143–156
16. Sternbach, R.A., Tursky, B.: Ethnic differences among housewives in psychophysical and skin potential responses to electric shock. *Psychophysiology* 1 (1965) 241–246
17. Kim, J., André, E.: Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (2008) 2067–2083
18. Llabre, M.M., Spitzer, S.B., Saab, P.G., Ironson, G.H., Schneiderman, N.: The reliability and specificity of delta versus residualized change as a measure of cardiovascular reactivity to behavioral challenges. *Psychophysiology* 28 (1991) 701–711
19. Boucsein, W.: *Electrodermal activity*. New York, NY, USA: Plenum Press (1992)
20. Fridlund, A.J., Cacioppo, J.T.: Guidelines for human electromyographic research. *Psychophysiology* 23 (1986) 567–589
21. Chaudhuri, S., Pawar, T.D., Duttagupta, S.: *Ambulation analysis in wearable ECG*. New York, NY, USA: Springer Science+Business Media (2009)