PREREQUISITES FOR AFFECTIVE SIGNAL PROCESSING (ASP) – PART III

Egon L. van den Broek

http://www.human-centeredcomputing.com/
vandenbroek@acm.org

Joris H. Janssen, Marjolein D. van der Zwaag

User Experience Group, Philips Research Europe, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
{joris.h.janssen,marjolein.van.der.zwaag}@philips.com

Jennifer A. Healey

Future Technology Research, Intel Labs Santa Clara, Juliette Lane SC12-319 Santa Clara CA 95054, USA
jennifer.healey@intel.com

Abstract: This is the third part in a series on prerequisites for affective signal processing (ASP). So far, six prerequisites were identified: validation (e.g., mapping of constructs on signals), triangulation, a physiology-driven approach, and contributions of the signal processing community (van den Broek et al., 2009) and identification of users and theoretical specification (van den Broek et al., 2010). Here, two additional prerequisites are identified: integration of biosignals, and physical characteristics.

1 INTRODUCTION

This paper is the third paper in a series that addresses prerequisites for Affective Signal Processing (ASP). Let us start with what Rosalind W. Picard states in the preface of her book *Affective Computing* (AC) (Picard, 1997, p. x):

But what I ran into, in trying to understand how our brains accomplish vision, was emotion. Not as a corollary, tacked on to how humans see, but as a direct component, an integral part of perception. . . . The role of emotions in "being emotional" is a small part of their story. The rest is largely untold, and has profound consequences – not just for understanding human thinking, but specifically for computing.

Subsequently, she continues denoting the importance of emotions through stating that . . . *emotions play an essential role in rational decision making, perception, learning, and a variety of other cognitive functions* (Picard, 1997, p. x). She even poses that . . . *too little emotion can impair decision making* (Picard, 1997, p. x). The former notion was already embraced by cognitive sciences and is now generally accepted. The latter notion was a shocking conclusion for the artificial intelligence (AI) community since their traditional foundation is one of reasoning and logic.

In a nutshell, AC aims to model or classify human emotions, using the 'affective signals' they transmit. These can be either facial characteristics (e.g., obtained through computer vision), movements, speech processing, biosignals, or a combination of these signals (van den Broek et al., 2009). This paper addresses the use of biosignals for affective signal processing (ASP), a notion which originates from the late 19th century (James, 1894). Biosignals are especially promising as they can now be measured unobtrusively and in real-time through wearable devices.

More than a decade after Picard's seminal book, more than anything else, it has become apparent that AC is incredibly hard and complex (Boehner et al., 2007; Chanel et al., 2009). Despite the vast efforts toward AC, results are still disappointing. Although it should be noted that some results, should be marked as good; e.g., Healey and Picard (1998); Rani et al. (2003); Hönig et al. (2007); Benovoy et al. (2008).

Although sometimes promising results are reported in literature, often follow-up research failed to replicate those results; cf. Chanel et al. (2009). Moreover, in the cases recognition rates of 90% or higher are achieved, three out of the four best results are based one the results on one participant. Only in the research of Hönig et al. (2007) a group of (24) participants participated; however, this study only differentiated between two levels of stress. Taken together, AC

has a long way to go, before large-scale real-world applications using AC are feasible.

Taken together, we pose that AC is a bridge too far and ASP is what should have our attention first. To enable a breakthrough in results on ASP, we propose to adopt a set of prerequisites, with which we started in van den Broek et al. (2009) and continued in van den Broek et al. (2010). After this, AC can be brought to practice.

This paper complements the two other papers van den Broek et al. (2009) and van den Broek et al. (2010), which already introduced 6 prerequisites. Together, the three papers should form the foundation for more successful ASP and, in the future, successful AC.

In the next section, the two new prerequisites will be introduced, namely: integration of biosignals and physiological characteristics. We end this paper in Section 3, with a brief conclusion.

2 PREREQUISITES – PART III

In van den Broek et al. (2009) and van den Broek et al. (2010), the following prerequisites for ASP were introduced: validity, triangulation, a physiology-driven approach, contributions from signal processing, user identification, and theoretical specification. While each of these is still of the utmost importance for ASP, we will now denote two additional prerequisites: physical characteristics and integration of biosignals.

2.1 Integration of biosignals

Although not frequently discussed, biosignals are influenced by other factors besides affect (Cacioppo and Tassinary, 1990). For instance, Figure 1 illustrates how pervasive motion artifacts can be for ASP in real world settings. Both HR and EDA are elevated during the period of high activity from 27 to 30 minutes. Moreover, the signal graphs also show that changes in HR follow changes in activity much more rapidly than EDA, in onset and especially in terms of recovery. For level 4 (walking) in this graph, it even seems that the physical effects are so dominant that ASP should not be attempted. In contrast, with level 1 (lying down), 2 (sitting), and 3 (standing/strolling) this is possible.

In van den Broek et al. (2009), we discussed dealing with these issues through triangulation; i.e., using multiple signals to describe or measure one construct. As a particular and highly effective instance of triangulation, we now discuss the integration of two or more biosignals into one feature that can be used as input to a classifier. This idea stems from the fact that additional biosignals can often explain noise that is present through other influences.

The integration of biosignals takes three steps: (1) identifying a theoretical relationship between multiple biosignals, (2) selecting an appropriate model that integrates both, and (3) data gathering and model training.

In the first step, a noisy biosignal is selected and theoretical relationships with other variables are identified. As we saw, HR is also influenced by physical activity and respiration. Hence, we should gather respiration data and accelerometer data to correct the HR signal for this noise. Other such relationships exist, for instance, between skin conductance and skin temperature, skin conductance and physical activity, or HR and skin temperature. Often this correction relates to a theoretical concept as well; e.g., correcting the high frequency (HF) power obtained from the inter beat intervals (IBI) for respiration gives a reliable measure for activity of the parasympathetic nervous system, which is involved in relaxation and recovery processes (Grossman and Taylor, 2007).

The second step consists of selecting a procedure to integrate multiple signals. A popular approach is using regression to describe the relationship between the variables. Correcting observed values then consists of computing its residualized value (i.e., the distance to the regression line) on the dimension of interest. In most cases, this is done using a linear regression line, but it can just as well be done with more complex non-linear relationships. A second popular method for the integration of such signals is through conditional probabilities. In that case, a Bayesian model is built in which the corrected value is conditional on the observed value and the values of the other influencing variables. When a new value is observed, the corrected value with the maximum a posteriori probability can be selected through this model.

In the third step, when the theoretical relationships and integration procedures are established, data has to be gathered that can be used to train the selected model. Each instance of the data should contain a value for of the features of the model. In the case of regression, there are many algorithms that determine a line or plane of best fit through the data. A particularly popular approach is using maximum likelihood approach by minimizing the least squares error. For Bayesian models, the data is modeled through (mostly continuous) probability distributions. The numerous methods for parameterizing a probability distribution from data are beyond the scope of this paper; see Bishop (2006) and Korb and Nicholson (2004) for more info.
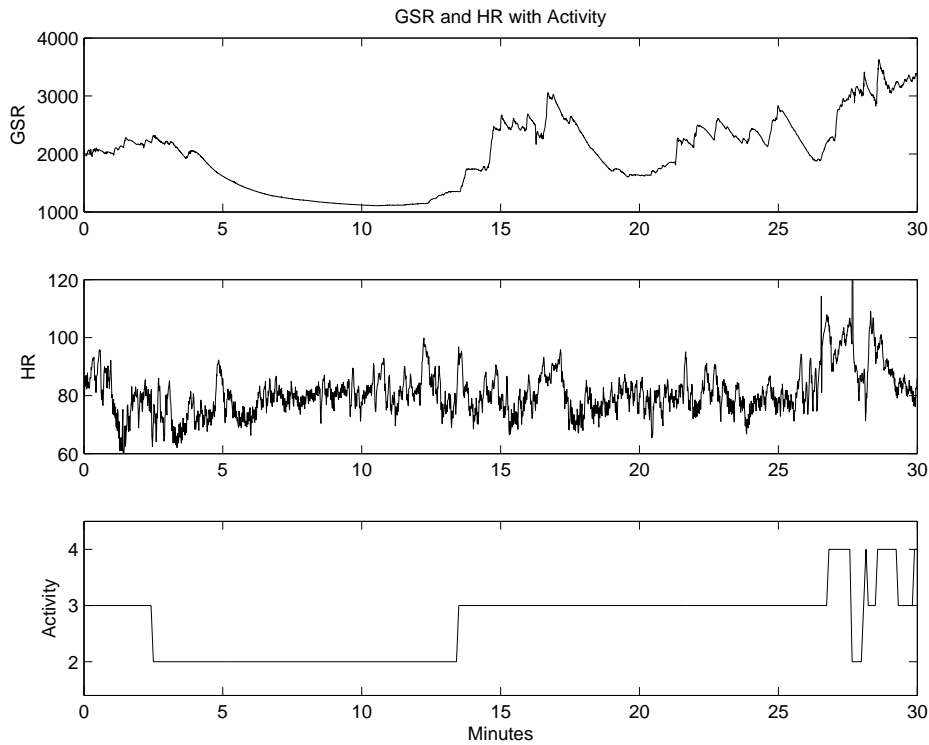
Figure 1: Recordings of Heart rate (HR), electrodermal activity (EDA) (or galvanic skin response, GSR), and a person's activity for a period of 30 minutes, in a real world setting. See Section 2.1 for the legend of the activities.

2.2   Physical characteristics

In this section, we discuss the implications of physical characteristics of sensors and the environment for affective signal processing. There are a number of different sensors. For respiration measurements, a gauge band can be placed around the chest. Thermistor sensors placed on the surface of the skin can be used to measure skin temperature (Kistler et al., 1998). HR can be measured through surface electrodes (ECG) or through a photoplethysmograph (BVP). Skin conductance and muscle tension (EMG) are also measured through surface electrodes.

The choice of surface electrodes depends on the kind of measurement, the aim of the measurement, and the application in which it is used. On the one hand, in the lab one opts for the most sensitive and reliable electrodes, which are wet electrodes that use a gel for better conductivity. On the other hand, for wearable affective measurements a better option is dry electrodes, as these are more practical and easier to attach and incorporate in devices.

The kind of gel used in wet electrodes depends on the measurement type. For skin conductance measurements, a salt less gel should be used as salt changes the composition of the skin which influences the measurement (Boucsein, 1992). For EMG and ECG, gels with high electric conductance are better.

The location of the surface electrodes is important as improperly placed can cause noise in the signal. However, in the case of ASP, the wearable devices and setting will put constraints on the location of the sensors. For example, the upper phalanx of the finger tips conventionally used for skin conductance measurements cannot be used while driving a car. Other parts of the hands or the sole of the foot should be used instead. For HR, instead of using electrodes on the chest (ECG) one can use a BVP sensor on the ear, hand, or foot. Skin temperature can also be measured on the foot instead of the hand.

For ASP, the physical characteristics of the environment like humidity and temperature also play an important role. This predominantly influences the skin conductance and temperature measurements. This point is of special interest for longer periods of continuous measurements and is also different in medical experiments which require a controlled lab situation in where humidity and temperature of the room can be kept constant.

To deal with the issues of different sensor positions and changes in environmental temperature and humidity one should standardize the measurements

using z-scores for each session. During continuous longer term measurements one can use a sliding time window for a set period, e.g. one or two hours, which is used for standardization.

To conclude, due to the large amount of differences in the aim of physiological measurements, different sensor positions, and different or even changing environmental conditions, one should always carefully puzzle to find the best combination of electrode types and locations. Furthermore, standardizing the signals will also reduce a lot of the otherwise unexplained variance in the signal. In the end, this will provide cleaner signals to the machine learning algorithms and will lead to a much more successful ASP.

3 CONCLUSION

This paper provided the third set of prerequisites for ASP. It comprises the prerequisites integration of biosignals and physical characteristics, which are complementary to the six previously introduced prerequisites: identification of users and theoretical specification (van den Broek et al., 2010) and validity, triangulation, the physiology-driven approach, and contributions of signal processing (van den Broek et al., 2009).

Perhaps the conclusion should be that, for now, AC is too complex (Boehner et al., 2007); cf. Chanel et al. (2009). We pose that it would be wise to take a step back, and study ASP, using the prerequisites provided. Then, time will learn whether AC will be future or remain fiction.

ACKNOWLEDGMENTS

REFERENCES

Benovoy, M., Cooperstock, J. R., and Deitcher, J. (2008). Biosignals analysis and its application in a performance setting: Towards the development of an emotional-imaging generator. In *Biosignals 2008: Proceedings of the first International Conference on Biomedical Electronics and Devices*, volume 1, pages 253–258, Funchal, Madeira, Portugal. INSTICC.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. New York, NY, USA: Springer Science+Business Media, LLC.

Boehner, K., DePaula, R., Dourish, P., and Sengers, P. (2007). How emotion is made and measured. *International Journal of Human-Computer Studies*, 65(4):275–291.

Boucsein, W. (1992). *Electrodermal activity*. New York, NY, USA: Plenum Press.

Cacioppo, J. and Tassinary, L. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45(1):16–28.

Chanel, G., Kierkels, J. J. M., Soleymani, M., and Pun, T. (2009). Short-term emotion assessment in a recall paradigm. *International Journal of Human-Computer Studies*, 67(8):607–627.

Grossman, P. and Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, 74(2):263–285.

Healey, J. A. and Picard, R. W. (1998). Digital processing of affective signals. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP98)*, volume 6, pages 3749–3752, Seattle, WA, USA. IEEE.

Hönig, F., Batliner, A., and Nöth, E. (2007). Real-time recognition of the affective user state with physiological signals. In *Proceedings of the Doctoral Consortium of Affective Computing and Intelligent Interaction (ACII)*, pages 1–8, Lisbon, Portugal.

James, W. (1894). The physical basis of emotion. *Psychological Review*, 1:526–529.

Kistler, A., Mariauzouls, C., and von Berlepsch, K. (1998). Fingertip temperature as an indicator for sympathetic responses. *International Journal of Psychophysiology*, 29(1):35–41.

Korb, K. B. and Nicholson, A. E. (2004). *Bayesian Artificial Intelligence*. Boca Raton, FL, USA: Chapman & Hall / CRC Press.

Picard, R. W. (1997). *Affective Computing*. Boston, MA, USA: MIT Press.

Rani, P., Sarkar, N., Smith, C. A., and Adams, J. A. (2003). Affective communication for implicit human-machine interaction. In *IEEE International Conference on Systems, Man, and Cybernetics*, volume 5, pages 4896–4903.

van den Broek, E. L., Janssen, J. H., van der Zwaag, M. D., and Healey, J. A. (2010). Prerequisits for Affective Signal Processing (ASP) – Part II. In *Biosignals 2010: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, page [submitted], Valencia – Spain.

van den Broek, E. L., Janssen, J. H., Westerink, J. H. D. M., and Healey, J. A. (2009). Prerequisits for Affective Signal Processing (ASP). In Encarnação, P. and Veloso, A., editors, *Biosignals 2009: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, pages 426–433, Porto – Portugal.