PREREQUISITES FOR AFFECTIVE SIGNAL PROCESSING (ASP) – PART II

Egon L. van den Broek
http://www.human-centeredcomputing.com/
vandenbroek@acm.org

Joris H. Janssen
User Experience Group, Philips Research Europe, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
joris.h.janssen@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

Marjolein D. van der Zwaag
User Experience Group, Philips Research Europe, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
marjolein.van.der.zwaag@philips.com

Keywords: affective signal processing, emotion, user identification, theoretical specification

Abstract: Last year, in van den Broek et al. (2009a), a start was made with defining prerequisites for affective signal processing (ASP). Four prerequisites were identified: validation (e.g., mapping of constructs on signals), triangulation, a physiology-driven approach, and contributions of the signal processing community. In parallel with this paper, in van den Broek et al. (2010) another set of two prerequisites is presented: integration of biosignals and physical characteristic. This paper continues this quest and defines two additional prerequisites: identification of users and theoretical specification. In addition, the second part of a review on the classification of emotions through ASP is presented; the first part can be found in van den Broek et al. (2009a).

1 INTRODUCTION

Almost half a century ago, Ulric Neisser (1963, p. 194) described *three fundamental and interrelated characteristics of human thought that are conspicuously absent from existing or contemplated computer programs.*

1. *Human thinking always takes place in, and contributes to, a cumulative process of growth and development.*
2. *Human thinking begins in an intimate association with emotions and feelings which is never entirely lost.*
3. *Almost all human activity, including thinking, serves not one but a multiplicity of motives at the same time.*

Nonetheless, artificial intelligence (AI) aimed at understanding human thought and developing computational and executable models of human thought *without* considering these three notions. Although, nowadays, a computer can beat the world's best chess players, the general opinion is that AI has failed. We, and others, think that Ulric Neisser's words are of vital importance and should be brought into AI practice.

In this paper, we will treat the second issue Neisser raised, that of emotions and feelings or, in other words, affect. Ever since Picard's book *Affective Computing* (AC) 1997, this direction of research has received growing interest. However, as with AI, the results with AC are disappointing. This is explained by the fact that research relevant for AC is scattered over a broad range of sciences and lacks generalization and robustness.

To force a breakthrough in results on AC we propose to consider a set of prerequisites for affective signal processing (ASP), before starting with AC in practice. The first part of these prerequisites was introduced last year in van den Broek et al. (2009a). This set of prerequisites was, however, not complete. This paper introduces the second part of the prerequisites on ASP. In parallel, the third part of the prerequisites on ASP is introduced in van den Broek et al. (2010). Together, these three papers should form the foundation for more successful ASP and AC.

In the next section, we will briefly denote ASP and AC, including a review presented in Table 1. This table contains the second part of our survey on ASP and AC, complementary to the part presented in van den Broek et al. (2009a). Section 3 introduces two new prerequisites for successful ASP, complementary to both those introduced in van den Broek et al. (2009a) and in van den Broek et al. (2010). Finally, we draw conclusions and denote the prerequisites' implications for applications on ASP.

2 AFFECTIVE SIGNAL PROCESSING (ASP)

ASP is often employed from three specialized areas of signal processing:

- movement analysis (Gunes and Piccardi, 2009),
- computer vision techniques (Gunes and Piccardi, 2009), and
- speech processing (Ververidis and Kotropoulos, 2006).

However, these signals still have their major disadvantages. In contrast, such issues have been resolved for biosignals in recent years: currently, high fidelity, cheap, and unobtrusive biosignal recordings are easy to obtain. Moreover, the recording devices can be easily integrated in various products (van den Broek and Westerink, 2009; Gamboa et al., 2009). Therefore, this paper focusses on biosignals. For an overview of the most commonly used biosignals and their features, we refer to van den Broek et al. (2009a).

The review in Table 1 illustrates both differences and similarities among research on AC, conducted over the last decade. As this table shows, most studies recorded people's cardiovascular and electrodermal activity. However, differences between the studies prevail over the similarities. The number of participants ranges from 1 to 50, although studies including > 20 participants are relatively rare; cf. Table 1. The number of features determined through ASP also ranges considerably; i.e., from 3 to 193. Only half of the studies applied feature selection/reduction, where this would be advisable in general.

For AC, a broad plethora of classifiers are used. The characteristics of the categories among which has to be discriminated is different from most other classification problems. The emotion classes used are typically ill defined, which makes it hard to compare studies. Moreover, the number of emotion categories (i.e., the classes) to be discriminated ranges considerably: from 2 to 8. Although these are small numbers in terms of pattern recognition and machine learning, the results are behind that of other classification problems. With AC recognition rates of 60%–80% are common, where in most other pattern recognition problems, recognition rates of $> 90\%$ and often $> 95\%$ are often reported. This illustrates the complex nature of AC and the need to consider prerequisites for ASP.

3 PREREQUISITES – PART II

In van den Broek et al. (2009a), the first set of prerequisites for ASP was introduced: validity, triangulation, a physiology-driven approach, and contributions from signal processing. One additional set of prerequisites is presented in van den Broek et al. (2010) and comprises: physical characteristics and integration of biosignals. Here we present a third set, which complements the two other, by discussing user identification and theoretical specification.

3.1 Tailored ASP – User identification

Throughout the field of AC, an ongoing debate is present on generic versus personal approaches to emotion recognition. Some research groups specialized in AC have moved from general AC to AC for specialized groups or individuals. For example, the group of Picard currently focusses on autism (Picard and Goodwin, 2008). In general, the identification of users has major implications for ASP. We propose three distinct categories, among which research in affective science could choose:

1. *all*: generic ASP; see also Table 1 and van den Broek et al. (2009c); van den Broek and Westerink (2009)
2. *group*: tailored ASP; e.g., Choi and Woo (2005); Sternbach and Tursky (1965)
3. *individual*: personalized ASP; e.g., Picard et al. (2001); Healey and Picard (2005)

Although attractive from a practical point of view, the category *all* will probably not solve the mysteries concerning affect. As is long known in neurology and psychology, special cases can help in improving ASP.

For the categories *group* and *individual*, the following subdivision can be made:

1. Specific characteristics; e.g., autism (Picard and Goodwin, 2008)
2. Psychological traits; e.g., Personality (Krohne et al., 2002; van den Broek et al., 2009c) or empathic intelligence (Håkansson and Montgomery, 2003).

Table 1: An overview of 18 studies on automatic biosignal-driven classification of emotions of the last decade.

information source	year	signals	participants	number of features	selection / reduction	classifiers	target	classification result
Picard et al.	2001	$\mathcal{C},\mathcal{E},\mathcal{R},\mathcal{M}$	1	40	SFS, Fisher	LDA	8 emotions	81%
Lisetti & Nasoz	2002	$\mathcal{C},\mathcal{E},\mathcal{S}$	10	3	-	kNN, LDA	5 emotions	85%*
Kim et al.	2002	$\mathcal{C},\mathcal{E},\mathcal{S}$	50	10	-	SVM	4 emotions	61%
			50/125	10	-	SVM	3 emotions	55% / 78%
Kim et al.	2004	$\mathcal{C},\mathcal{E},\mathcal{S}$	50	10	Fisher	SVM	4 emotions	62%
							3 emotions	78%
Rani et al.	2004	$\mathcal{C},\mathcal{E},\mathcal{M}$	1	6	-	FLS	2 anxiety levels	??%
Healey & Picard	2005	$\mathcal{C},\mathcal{E},\mathcal{R},\mathcal{M}$	9	22	Fisher	LDA	3 stress levels	97%
Kim et al.	2005	$\mathcal{C},\mathcal{E},\mathcal{R},\mathcal{M}$	3	26	SFS	LDF	4 emotions	53% / 74%
Lisetti & Nasoz	2005	$\mathcal{C},\mathcal{E},\mathcal{S}$	41	86	-	kNN, ANN (2x)	2 emotions (3 sets of)	92%
Liu et al.	2005	$\mathcal{C},\mathcal{E},\mathcal{M}$	15	13(?)	-	kNN, RT, BN, SVM	5 emotions	86%
Liu et al.	2006	$\mathcal{C},\mathcal{E},\mathcal{M},\mathcal{S}$	14	35	-	RT	3 anxiety levels	70%
Rainville et al.	2006	$\mathcal{C},\mathcal{G},\mathcal{S},\mathcal{M},\mathcal{P}$	43	18	PCA	LDA	4 emotions	49%
Jones & Troen	2007	$\mathcal{C},\mathcal{E},\mathcal{R}$	13	11	-	ANN	5 arousal levels	31% / 62%
							5 valence levels	26% / 57%
Yang & Liu	2007	$\mathcal{C},\mathcal{E},\mathcal{R},\mathcal{M}$	1	193	BPSO	kNN	4 emotions	86%
Kim	2007	$\mathcal{C},\mathcal{E},\mathcal{R},\mathcal{M},\mathcal{S}$	3	77	SBS	kNN, ANN, LDA	4 emotions	51%–71%
Cheng & Liu	2008	$\mathcal{M}$	1	14	DWT	ANN	4 emotions	75%
Lichtenstein et al.	2008	$\mathcal{C},\mathcal{E},\mathcal{R},\mathcal{M},\mathcal{S}$	40	5	-	SVM	5 emotions	47%
							2 levels of arousal	82%
							2 levels of valence	72%
Chanel et al.	2009	$\mathcal{C},\mathcal{E},\mathcal{R}$	10	18	-	LDA, SVM, RVM	3 emotions	51%
							2 emotions	66%
Van den Broek et al.	*in press*	$\mathcal{E},\mathcal{M}$	21	10	ANOVA, PCA	kNN, SVM, ANN	4 emotions	61%

Signals: $\mathcal{C}$: cardiovascular activity; $\mathcal{E}$: electrodermal activity; $\mathcal{R}$: respiration; $\mathcal{M}$: electromyogram; $\mathcal{S}$: skin temperature; and $\mathcal{P}$: pupil diameter. *Selection:* PCA: Principal Component Analysis; SFS: Sequential Forward Selection; SBS: Sequential Backward Selection; BPSO: Binary Particle Swarm Optimization; DWT: Discrete Wavelet Transform; Fisher: Fisher projection; and ANOVA: ANalysis Of VAriance. *Classifiers:* RT: Regression Tree; BN: Bayesian Network; ANN: Artificial Neural Network; SVM: Support Vector Machine; RVM: Relevance Vector Machines; LDA: Linear Discriminant Analysis; kNN: k-Nearest Neighbors; and FLS: Fuzzy Logic System. *Note.* * The authors present 100% successful classification for two emotion categories. However, this might indicate a questionable training and testing setup.

3. Demographics; e.g., age, sex, race (Sternbach and Tursky, 1965), or level of education (van den Broek et al., 2009c).
4. Activities; e.g., office work (Janssen et al., 2009), driving a car (Healey and Picard, 2005) or flying a plane, and running (Healey, 2009).

This subdivision is based current practice with ASP; however, possibly it should be altered.

So far, comparisons between research results on ASP are mostly made between results of either individuals or groups selected to resemble the general population; cf. Table 1. However, user-tailored approaches should be explored as well. In particular, experiences with specific groups can substantially contribute to the further development of ASP, as has been seen in other sciences; e.g., biology, psychology, and medicine. But also, individual biosignal response patterns should be taken into consideration, since people's affective signals differ widely.

Having said that, the question remains, how to handle this striking variety between people. We present three approaches, which are on another level than usually adopted, but can tackle these problems:

Hybrid classification systems (Berzal et al., 2004). Most often, such architectures incorporate both a (logic-based) reasoning system and a machine learning component. To the authors knowledge, so far, this approach has not been applied for ASP. It has, however, been applied successfully for speech-based emotion recognition (Schuller et al., 2004).

Multi-agent systems and multi-classifier systems. Two approaches within this field could be of interest: 1) Multi-layered architectures, where each layer determines the possible classes to be processed or the classifiers to be chosen for the next layer and 2) An ensemble of classifiers, trained on the same or distinct biosignals and their features. Their outputs are collected into one compound classification, often determined through a voting scheme. For more information on this topic, we refer to Lam and Suen (1995) and Kuncheva (2002).

Biosignal signatures. Related to schemes that are used in forensics (Rogers, 2003), ASP could benefit from personalized profiles or schemes that tailor to a generic profile to people's unique biosignal signatures. Moreover, this approach could be extended to incorporate context information, as is already done in forensics (Rogers, 2003). Biosignal signatures require advanced multi-modal data mining and knowledge discovery strategies, and is related to the baseline matrix as proposed by Picard et al. (2001).

Each of these approaches enable processing of multi-modal data, which allows to incorporate a range of characteristics. This makes them promising for ASP applications, also outside the scope of user identification.

3.2 Theoretical specification

Changes in biosignals relate to changes in many psychological constructs Cacioppo and Tassinary (1990). For ASP, it is important to distinguish between these different psychological constructs. This involves two different situations: Firstly, biosignals can have more or less equal response patterns but in different time frames; e.g., short time frames for emotions and longer time frames for moods. Second, biosignals can have the same response patterns in the same time frames but still relate to different psychological constructs; an increase in inspiration rate can imply increased positive moods but also increased task demands or mental effort. We propose three ways of dealing with this complexity: (1) specification of the relation between construct of interest and biosignal, (2) involving context information, and (3) using multiple classifier systems.

In the first place, a thorough description of the relation between the construct of interest and the biosignals is necessary. By doing this, distinguishing biosignal properties for the construct of interest can be found. For instance, when classifying emotions short-term changes are of interest, whereas when classifying moods only long-term changes are relevant. In addition, when trying to distinguish workload from mood, one should not be interested in skin conductance changes. Instead, one could look at heart rate variability as this typically reacts stronger to workload than mood.

Second, context provides a lot of information on the psychological constructs which might have been changed. For example, while driving a car workload is changing quickly depending on the road situation while your mood is likely to remain equal. On the other hand, during watching a television show changes in biosignals are more likely to come from affect induction than from changes in cognitive load, motivation, or memory. By inserting context information, e.g. captured by a camera, the probability that changes will occur in specific psychological constructs can be modeled into the system. Thereby, increasing the change of allocating changes in biosignals to changes in the correct psychological construct.

Finally, one can use multiple classifiers to make a classification of all separate psychological constructs. In turn, the construct that with the most certain classification can be selected as the influenced construct. Moreover, an extra classifier can be trained that receives it's input from the separate classifiers and makes a selection based on this information.

To conclude, these three ways can help to deal with the problem of the many-to-many relationship between psychology and physiology. Note that we have assumed that the psychological constructs are independent of each other, which is actually not the case. Nonetheless, treating them as if being independent of each other is necessary for practical purposes.

4 CONCLUSION

This paper provided the second part of prerequisites for ASP. For the first and third part, we refer to respectively van den Broek et al. (2009a) and van den Broek et al. (2010). Here, the prerequisites identification of users and theoretical specification are introduced. These prerequisites are complementary to those presented in the other two papers: validity, triangulation, the physiology-driven approach, and contributions of signal processing (van den Broek et al., 2009a) and physical characteristics and integration of biosignals (van den Broek et al., 2010). Moreover, the second part of a review on ASP has been presented; see Table 1, complementary to the review table presented in van den Broek et al. (2009a).

The review (see Table 1) and the prerequisites, illustrate the complexity and lack of success of AC. This urges us to emphasize that a step back should be made by looking at prerequisites for successful ASP to achieve true progress on a later stage, instead of running forward and ignoring the problems encountered in previous studies. We sincerely hope that the prerequisites can contribute to or even guide the promising future ASP provides us with.

ACKNOWLEDGMENTS

The authors would like to thank Joyce H.D.M. Westerink (Philips Research, Eindhoven, The Netherlands) for her comments on an earlier versions of this paper. Furthermore, we would like to thank the anonymous reviewers, who provided us the opportunity to improve this paper.

REFERENCES

Berzal, F., Cubero, J.-C., Sánchez, D., and Serrano, J. M. (2004). ART: A hybrid classification model. *Machine Learning*, 54(1):67–.

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.

Cheng, B. and Liu, G. (2008). Emotion recognition from surface EMG signal using wavelet transform and neural network. In *Proceedings of The 2nd International Conference on Bioinformatics and Biomedical Engineering (ICBBE)*, pages 1363–1366, Shanghai, China. Piscataway, NJ, USA: IEEE Press.

Choi, A. and Woo, W. (2005). Physiological sensing and feature extraction for emotion recognition by exploiting acupuncture spots. *Lecture Notes in Computer Science (Affective Computing and Intelligent Interaction)*, 3784:590–597.

Gamboa, H., Silva, F., Silva, H., and Falcão, R. (2009). PLUX – Biosignals Acquisition and Processing. URL: `http://www.plux.info` [Last accessed on November 02, 2009].

Gunes, H. and Piccardi, M. (2009). Automatic temporal segment detection and affect recognition from face and body display. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, 39(1):64–84.

Håkansson, J. and Montgomery, H. (2003). Empathy as an interpersonal phenomenon. *Journal of Social and Personal Relationships*, 20(3):267–284.

Healey, J. A. (2009). Affect detection in the real world: Recording and processing physiological signals. In *Proceedings of the IEEE 3rd International Conference on Affective Computing and Intelligent Interaction, ACII*, volume 1, pages 729–734, Amsterdam, The Netherlands. IEEE Press.

Healey, J. A. and Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2):156–166.

Janssen, J. H., van den Broek, E. L., and Westerink, J. H. D. M. (2009). Personalized affective music player. In *Proceedings of the IEEE 3rd International Conference on Affective Computing and Intelligent Interaction, ACII*, volume 1, pages 472–477, Amsterdam, The Netherlands. IEEE Press.

Jones, C. M. and Troen, T. (2007). Biometric valence and arousal recognition. In Thomas, B. H., editor, *Proceedings of the Australasian Computer-Human Interaction Conference (OzCHI)*, pages 191–194, Adelaide, Australia.

Kim, J. (2007). *Bimodal emotion recognition using speech and physiological changes*, chapter 15, pages 265–280. Vienna, Austria: I-Tech Education and Publishing.

Kim, J., André, E., Rehm, M., Vogt, T., and Wagner, J. (2005). Integrating information from speech and physiological signals to achieve emotional sensitivity. In *Proceedings of the 9th European Conference on Speech Communication and Technology*, pages 809–812, Lisboa, Portugal. Lisboa, Portugal: L2F – Spoken Language Systems Laboratory.

Kim, K. H., Bang, S. W., and Kim, S. R. (2002). Development of person-independent emotion recognition system based on multiple physiological signals. In *Proceedings of the Second Joint EMBS–BMES Conference*, volume 1: Bioengineering – Integrative Methodologies, New Technologies, pages 50–51, Houston, Texas, USA.

Kim, K. H., Bang, S. W., and Kim, S. R. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical & Biological Engineering & Computing*, 42(3):419–427.

Krohne, H. W., Pieper, M., Knoll, N., and Breimer, N. (2002). The cognitive regulation of emotions: The role of success versus failure experience and coping dispositions. *Cognition & Emotion*, 16(2):217–243.

Kuncheva, L. I. (2002). A theoretical study on six classifier fusion techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):281–286.

Lam, L. and Suen, C. Y. (1995). Optimal combinations of pattern classifiers. *Pattern Recognition Letters*, 16(9):945–954.

Lichtenstein, A., Oehme, A., Kupschick, S., and Jürgensohn, T. (2008). Comparing two emotion models for deriving affective states from physiological data. *Lecture Notes in Computer Science (Affect and Emotion in Human-Computer Interaction)*, 4868:35–50.

Lisetti, C. L. and Nasoz, F. (2002). MAUI: a multimodal affective user interface. In *Proceedings of the tenth ACM international conference on Multimedia*, pages 161–170, Juan les Pins, France. New York, NY, USA: ACM Press.

Lisetti, C. L. and Nasoz, F. (2005). Affective intelligent car interfaces with emotion recognition. In *Proceedings of 11th International Conference on Human-Computer Interaction*, pages [on cd–rom], Las Vegas, Nevada, USA. Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc.

Liu, C., Rani, P., and Sarkar, N. (2005). An empirical study of machine learning techniques for affect recognition in human-robot interaction. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2662–2667, Edmonton, Alberta, Canada. Piscataway, NJ, USA: IEEE Press.

Liu, C., Rani, P., and Sarkar, N. (2006). Human-robot interaction using affective cues. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06)*, pages 285–290, Hatfield, UK. IEEE Computer Society.

Neisser, U. (1963). The imitation of man by machine – the view that machines will think as man does reveals misunderstanding of the nature of human thought. *Science*, 139(3551):193–197.

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

Picard, R. W. and Goodwin, M. (2008). Developing innovative technology for future personalized autism research and treatment. *Autism Advocate*, 50(1):32–39.

Picard, R. W., Vyzas, E., and Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1175–1191.

Rainville, P., Bechara, A., Naqvi, N., and Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, 61(1):5–18.

Rani, P., Sarkar, N., Smith, C., and Kirby, L. (2004). Anxiety detecting robotic systems - towards implicit human-robot collaboration. *Robotica*, 22(1):85–95.

Rogers, M. (2003). The role of criminal profiling in the computer forensics process. *Computers & Security*, 22(4):292–298.

Schuller, B., Rigoll, G., and Lang, M. (2004). Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine – belief network architecture. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, volume 1, pages I–577–580, Montreal, QC, Canada.

Sternbach, R. A. and Tursky, B. (1965). Ethnic differences among housewives in psychophysical and skin potential responses to electric shock. *Psychophysiology*, 1(3):241–246.

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 III. In *Biosignals 2010: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, page [this volume], Valencia – Spain.

van den Broek, E. L., Janssen, J. H., Westerink, J. H. D. M., and Healey, J. A. (2009a). 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.

van den Broek, E. L., Lisý, V., Janssen, J. H., Westerink, J. H. D. M., Schut, M. H., and Tuinenbreijer, K. (2009b). *Affective Man-Machine Interface: Unveiling human emotions through biosignals*, page [in press]. Biomedical Engineering Systems and Technologies: BIOSTEC2009 Selected Revised papers. Berlin/Heidelberg, Germany: Springer.

van den Broek, E. L., Schut, M. H., Westerink, J. H. D. M., and Tuinenbreijer, K. (2009c). Unobtrusive Sensing of Emotions (USE). *Journal of Ambient Intelligence and Smart Environments*, 1(3):287–299.

van den Broek, E. L. and Westerink, J. H. D. M. (2009). Considerations for emotion-aware consumer products. *Applied Ergonomics*, 40(6):1055–1064.

Ververidis, D. and Kotropoulos, C. (2006). Emotional speech recognition: Resources, features, and methods. *Speech Communication*, 48(9):1162–1181.

Yang, Y.-H., Lin, Y.-C., Su, Y.-F., and Chen, H. H. (2007). Music emotion classification: A regression approach. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pages 208–211, Beijing, China. Piscataway, NJ, USA: IEEE, Inc.