Personalized affective music player

Joris H. Janssen
Eindhoven University of Technology
Philips Research
Eindhoven, The Netherlands
j.h.janssen@tue.nl

Egon L. van den Broek
University of Twente
Enschede, The Netherlands
vandenbroek@acm.org

Joyce H.D.M. Westerink
Philips Research
Eindhoven, The Netherlands
joyce.westerink@philips.com

Abstract

We introduce and test an affective music player (AMP) that selects music for mood enhancement. Through a concise overview of content, construct, and ecological validity, we elaborate five considerations that form the foundation of the AMP. Based on these considerations, computational models are developed, using regression and kernel density estimation. We show how these models can be used for music selection and how they can be extended to fit in other systems. Subsequently, the success of the models is illustrated with a user test. The AMP augments music listening, where its techniques, in general, enable automated affect guidance. Finally, we argue that our AMP is readily applicable to real-world situations as it can 1) cope with noisy situations, 2) handle the large inter-individual differences apparent in the musical domain, and 3) integrate context or other information, all in real-time.

1. Introduction

Whether it is at sporting events, while studying for an exam, at parties, or with religious rituals, music is intrinsically interwined with people's every-day life [37]. Music is an excellent mood inducer, both in laboratory settings [43] and in the real world [38]. Moreover, it has been suggested that mood enhancement is one of the most important functions of music [21, 22]. So, would it not be great if your music player knew what effects a song has on your mood? Such a music player could create playlists that energize you, relax you, or help you concentrate. Moreover, with some additional information about the situation you are in, it could even automatically select appropriate music. For instance, when coming home after work, it could select relaxing music, or when getting up, it could select music based on increasing energy levels to get you going.

The idea of an affective music player (AMP) is not new; e.g., [12, 26]. For instance, The Body Rest system of Liljedahl et al. [18] is a biofeedback system that works by altering tempo of the music based on heart rate (HR). The PAPA system of Oliver et al. [23] measures and saves HR responses to music, and uses this information to select music. The Affective Remixer of Chung et al. [6] uses skin conductance level (SCL) and foot tapping to infer affective responses to music, and rearranges music segments to induce a target state. All these systems use physiological signals that are known to be closely coupled to emotional changes.

Validation is an essential part of the development of any system. A post-hoc analysis, however, suggests that previous AMPs might lack content, construct, and ecological validition [41]. Lack of content validity is shown by weak relations between the physiological measures and the affective states represented by the music player. Lack of construct validity refers to a missing theoretical framework around the affective states and their relations with physiological changes. Finally, lack of ecological validity is evidenced by the data gathering being done in well-controled experimental sessions, instead of in the real world. Hence, it is unknown to what extent these music players will be able to handle all the noise occuring in more realistic situations.

In this paper, we take on the challenge of designing an AMP working in real-time. We start by elaborating five considerations from the psychological literature that form the design rationale of the music player. Next, we present the technical implementation of these considerations. Subsequenlty, we describe the results of a user test. Finally, we

978-1-4244-4799-2/09/$25.00 ©2009 IEEE

discuss the advantages and disadvantages of our approach and give suggestions for further research.

2. Design considerations

Research on affective technology is shaped by an inherently multidisciplinary range of fields. Therefore, it is eminent to take into account theory and methods from various disciplines. In the design of our AMP, we have identified five crucial considerations, constituting the design rationale of our AMP.

First, musical taste is highly personal. Accordingly, a number of personal factors can be identified that influence the effect of music on affect; e.g., personality [27], familiarity with the music [29, 34], age [24], gender [42], and musical preference [32, 34]. In line with this, it has been shown that personally selected music is a stronger affect inducer than experimenter-selected music [28]. Therefore, models for predicting the musical effects must be personalized (see also [16]).

Second, although affect is the common denominator of the terms emotion and mood [30], there are important differences between the two. In contrast to emotions, moods are (1) long lasting and change gradually [8], (2) not object related [10], and (3) often experienced without concurrent awareness of their origin [15]. The goal of the AMP is to direct one's affective state for a longer period than that of emotions; hence, we focus on moods. Moods change gradually and have effects of at least a number of minutes [30]. Therefore, we focus on effects over one or multiple songs, instead of within a song. Finally, moods are, like emotions, accompanied by physiological changes [20, 35]. These physiological changes can be measured unobtrusively in a wearable device and are, therefore, suited for an AMP.

Third, the affective loop, apparent in most physiological computing applications, can be defined through three steps: (a) infer user's current affective state from physiology, (b) set an affective goal state and select music based on the goal and current state, and (c) measure the affective physiological changes [1, 9]. However, the exact relation between affect and physiology is problematic [17, 25]. Moreover, machine learning studies trying to recognize affect from physiology report disappointing performances (as reviewed in [41]). In fact, it is a common critique on affective computing that automated recognition of affect is difficult, if not impossible [3, 14, 39]. Therefore, we do not try to automatically infer the affective state from the physiology. Instead, music selection is based on physiological changes and a physiological goal state. The physiological goal state can be inferred from psychological studies as pointed out in the previous paragraph; e.g., increasing SCL will increase arousal and increasing skin temperature (ST) will increase valence [28]. This overcomes the problematic inference step from physiology to affect but still allows to regulate mood in an affective physiological loop.

Fourth, physiology is responsive to many psychological and physical influences beside affect [5]. For instance, when standing up from a chair, HR increases. Similarly, there are many other factors that contaminate the affective information in the physiological signals. As we train and validate the AMP in a real-world setting, all kinds of uncontrollable factors will influence the physiological signals. We will deal with this in two ways. First, we use probabilities to model the physiological effects, which naturally deal with noisy data [2]. Second, we standardize the physiological signals per measurement session, to be able to compare different sessions.

Fifth and final, 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 [36, 44]. For instance, when ST is high, it tends to decrease. In turn, a stimulus that normally increases ST, might now keep it at the same high level. Its effect is cancelled out by the tendency of high ST to decrease. Thus, although the stimulus has an effect, this may be masked by consequences of the initial value. Therefore, we incorporate the principle of initial values in our models.

These five considerations form the foundation for the AMP. In the next Section, it will be described how these are implemented in our system.

3. System description

With the five considerations in mind, as were just introduced, we chose to model the effects of music in a probablistic manner. These models are trained during the course of system use; i.e., every time a song is listened to, the systems learns from the resulting physiological changes. How this is achieved will be described in the following sections.

3.1. Data preprocessing

First of all, every time a song k is listened to, we want to extract the outcome of that song k over the entire measurement session n (i.e., typically the entire day). The means of the physiological signals in the last minute of the song, denoted by x_{kn}, are extracted. Next, these means are standardized over that session n, using

$$z_{kn} = \frac{x_{kn} - \mu_n}{\sigma_n}, \tag{1}$$

where z_{kn} is the standardized value, μ_n is the mean and σ_n is the standard deviation over session n. This method has proven to be succesful for various physiological signals [4].

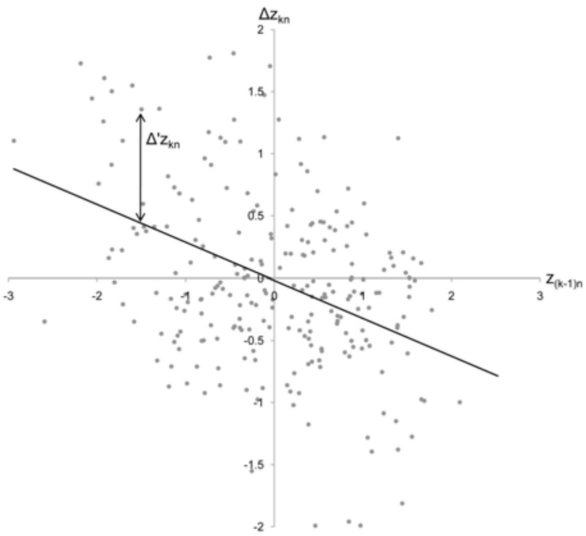


Figure 1. A regression line modelling the prestimulus-to-delta relationship. The gray dots represent measured datapoints, and the black line depicts the regression line.

Finally, delta scores Δz_{kn} are computed. These indicate the effect of the song k in session n on the physiology [19], by using

$$\Delta z_{kn} = z_{kn} - z_{(k-1)n}, \tag{2}$$

where $z_{(k-1)n}$ is the presong level, with $k \geq 2$.

3.2. Principle of initial values

As elaborated in the Introduction and defined in Equation 2, the delta score Δz_{kn} of a stimulus depends on the prestimulus level $z_{(k-1)n}$ [36]. Hence, we need to reduce the noise this effect generates for the variable we are interested in; namely, the physiological change the song typically elicits.
A regression line can be used to model this relation:

$$y\left(z\right) = w_1 z + w_0, \tag{3}$$

where w_0 and w_1 are the parameters of the regression line. After these parameters have been assessed, corrected delta scores $\Delta' z_{kn}$ are computed by subtracting the value of regression line y_{kn} at $z_{(k-1)n}$ from the delta scores Δz_{kn} (see also Figure 1):

$$\Delta' z_{kn} = \Delta z_{kn} - y\left(z_{(k-1)n}\right). \tag{4}$$

Note that these regression lines differ per person. Thus, it is important to estimate this relation for every person separately.

3.3. Personal probability distributions

To completely describe a random variable, the probability density function (pdf) can be applied. This positive real-valued function integrates to 1 and can be used to exploit important characteristics of the variable; e.g., mean, quantiles, or power spectral density [13]. Moreover, the pdf naturally deals with uncertainty in the data and fits the Bayesian paradigm [2]. Hence, we describe the physiological change a song elicits by a pdf over $\Delta' z$; i.e., the dimension on which corrected delta scores $\Delta' z_{kn}$ lie. However, the pdfs are unknown and we only have a limited number of observations of $\Delta' z$.

A well-established approach for estimating a pdf over observations is kernel density estimation (KDE) [33]. KDEs are unsupervised and nonparametric; i.e., they make no a priori assumptions of the underlying distribution and they can approximate any distribution. In addition, KDEs are asymptotically unbiased; i.e., they are unbiased when the sample tends to infinity. Moreover, the more sample points, the better the estimation of the pdf.

For every song k, the KDE puts a radial kernel function $\mathrm{K}\left(\Delta' z \mid \Delta' z_{kn}, h_k\right)$ with precision h_k around all N_k measured points $\Delta' z_{kn}$ of this song, as measured in the various sessions, and averages over all these N_k kernels:

$$p_k(\Delta' z) = \frac{1}{N_k} \sum_{n=1}^{N_k} \mathrm{K}\left(\Delta' z \mid \Delta' z_{kn}, h_k\right). \tag{5}$$

This yields the pdf $p_k(\Delta' z)$ over $\Delta' z$ for song k.

The selection of h is important: when h is chosen too large, the probability distribution is over smoothed; yet, when h is chosen too small, the probability distribution is under smoothed. Various methods have been proposed to calculate an accurate h; see [40] for a review. Because it is computationally efficient and robust against outliers, we propose to use

$$h_k = 1.06 \cdot \min\left(SD_k, \frac{R_k}{1.34}\right) \cdot N_k^{-\frac{1}{5}}, \tag{6}$$

where R_k is the interquartile range and SD_k is standard deviation of the corrected delta scores $\Delta' z_{kn}$ of song k; see [11] for details. In addition, a radial Gaussian kernel can be used, where the mean is $\Delta' z_{kn}$ and the standard deviation is h_k. This is often employed for its nice analytical properties. Moreover, it provides a smooth distribution [13]:

$$\mathrm{K}\left(\Delta' z \mid \Delta' z_{kn}, h_k\right) = \frac{1}{\sqrt{2\pi h_k^2}} \exp\left(-\frac{(\Delta' z - \Delta' z_{kn})^2}{2h_k^2}\right). \tag{7}$$

Taken together, the effects of the music can be modeled by KDEs through repeated exposure to the music. In the next section, we will discuss how these KDEs can be used to select music.

3.4. Music selection

Music can be selected based on these models in at least two ways. First, music can be selected to direct to one of the extremes of the physiological signal. For most purposes, this might be sufficient. For instance, when relaxing music is needed, music that lowers SCL should be selected as SCL is known to be highly correlated with arousal or excitement. Music selection, then, entails calculating the probability of each song in the increase $[0, \infty\rangle$ or decrease $\langle -\infty, 0]$ range and selecting the song with the highest probability. This mechanism can be extended to include a neutral state beside the two extremes, by including a neutral range c_0 around 0; e.g., $c_0 = 0.5$. Here, c_0 is a constant ≥ 0 that indicates the size of the neutral state.

Sometimes the optimal affective state is not reached at either one of the extremes of a physiological measure. For instance, when trying to focus on a task, too relaxing music will lower the working spirit, whereas too arousing music might distract one from the task [7]. So, in contrast to a general direction, the physiological models should also allow the direction of the physiological state of a listener towards a specific point. Directing to a specific point, however, is more difficult than steering in a general direction. It requires a wide variety of very specific music. Furthermore, the optimal physiological state or the physiological flow state has to be identified. As this state is probably highly personal and depends on the task and skills of the listener, this is problematic. Nonetheless, when this specific physiological state is known, the KDEs allow to select a song with the highest probability in a specific confidence interval c. The confidence interval c can be used to calculate a range around the point $\Delta' z_t$ one wants to direct to, using

$$\left[\frac{\Delta' z_t - c}{2}, \frac{\Delta' z_t + c}{2}\right].$$

Song selecting, then, entails selecting the song with the highest probability in this range.

4. User test

To test the proposed models, we measured physiological responses (SCL and ST) of three participants listening to their own music. For ecological validity, this was done during regular working activities (i.e., desk work) over multiple days for a total of 23 hours per participant.

After we obtained 9 datapoints per participant for several songs, we trained the AMP as indicated in the previous Section. In Figure 2 examples of the resulting KDEs are shown. Subsequently, we selected the songs with highest probabilities for increasing and decreasing ST, based on the procedure described in Section 3.4 ($c_0 = 0.5$). We asked the participants to express how these songs make them feel on a 5-point Likert scale ranging from sad (-2) to happy (2).

The ratings of the songs selected for the increasing range were compared to the ratings of the songs selected for the decreasing range. This was done through an independent samples t-test on the feeling ratings, with direction (positive / negative) as independent variable.

As expected, the songs in the negative ST direction ($m = 0.8$) were more positively rated than the songs in the positive ST direction ($m = -0.4$; $t(28) = 2.05$, $p < .05$), as high ST is related to low valence [31]. This shows that our models can indeed be succesfully employed to select music that directs to a certain valence level.

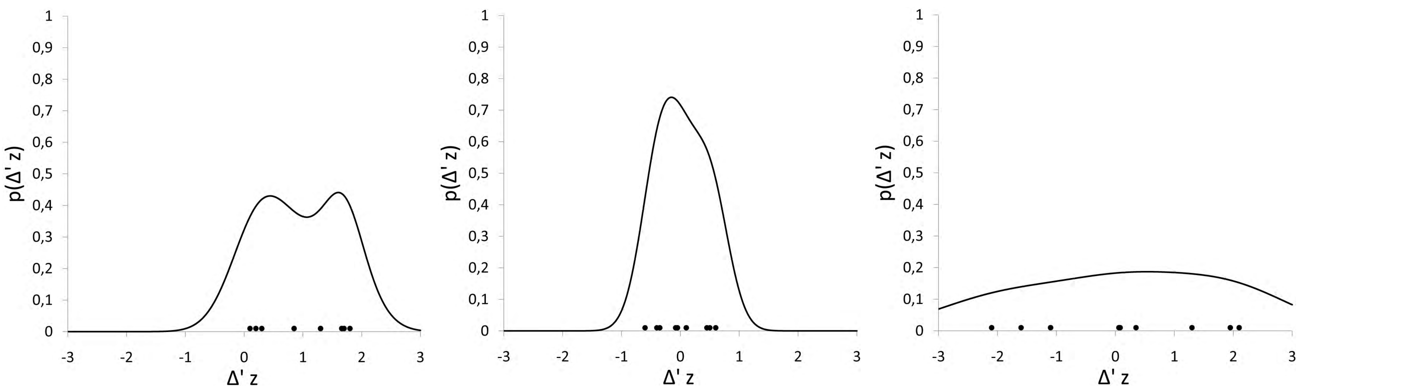
5. Discussion

Based on the five considerations elaborated in Section 2, we have shown how to design, implement, and use an AMP. Our AMP uses personalized probablistic models that naturally deal inter-individual and environmental noise. Moreover, we employ the principle of initial values in a physiological closed-loop system. Finally, instead of measuring short-term emotions, we focus on the longer-term effects of moods. In the next paragraphs, we will evaluate the design of the AMP and give directions for further research on some open issues.

In the introduction, we stated that the AMP should have high content, construct, and ecological validity. With content validity, we refer to the degree to which internal representations of the system are in line with the effects the music has. In our case, this means that the physiological representations can be used to influence the physiological state of the user. This is highly likely, as the physiological representations are the direct result of previous effects of the songs and there is no reason to assume that this would change.

With construct validity, we refer to the use of an affective theoretical framework around the physiological states. Here, the construct validity is high, as we have included temporal aspects of the affective and physiological changes in our models; see Section 2. In addition, our models are able to handle the noise generated by the principle of initial values. Finally, our empirical evidence supports the notion that directing based on physiology influences the affective state of the listener.

Ecological validity refers to the performance the system has in real-world situations, instead of lab settings. Through probablistic modelling, our music player naturally deals with noise in the environment. Moreover, our models are able to deal with the large inter-individual differences that are very apparent in the music domain. Finally, our models can easily be extended to bayesian networks. This way, other factors that influence the physiological changes (like context factors) can be included to increase ecological validity and the power of the predictions even more. Finally, our empirical test was done in a real-life setting, making the



(a) Increase (b) Neutral (c) No direction

Figure 2. Typical examples of KDEs. The dots depict measured values, and the line represents the KDE over the measured values. Figure (a) is an example of a song that increases the physiological level. Figure (b) has a neutral effect on physiological level. Figure (c) has no specific physiological effect, and can, thus, not be used by the AMP.

obtained evidence ecologically valid.

Although the models proposed are promissing, there are also some open issues. First of all, our current models can only be used to direct one physiological dimension. If we, however, consider the physiological changes to be independent, we can simply make selections based on multiple pdfs to direct several physiological signals simultaneously. For instance, select music that has a high probability of increasing SCL and decreasing heart rate. In case we consider the physiological changes to be correlated, which is likely so, multi-dimensional KDEs can be employed to model the effects of multiple physiological variables in one pdf. Second, the disadvantage of these models is that they require personal data before actual music selection can take place. This can be overcome by estimating prior distributions, based on music characteristics, personality, and music preference.

To conclude, through three levels of validation, we have shown how to succesfully develop an AMP. Five considerations from the psychological literature provided a solid starting point for our music player. As a result, our AMP works under noisy conditions, handles interpersonal differences, and allows the integration context information. This was supported by an extensive user test in a real-world situation. Moreover, it shows that a next generation of music players is within reach of commercial exploitation. This opens up various opportunities for automated affect direction with music, allowing affective technology to come to fruition in consumer products.

Acknowledgements

We would like to thank Marjolein van der Zwaag, Ard Biesheuvel, Ad Denissen, Tim Tijs, and Gert-Jan de Vries (Philips Research).

References

[1] J. Allanson and S. H. Fairclough. A research agenda for physiological computing. *Interacting with Computers*, 16:857–878, 2004.

[2] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, 2006.

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

[4] W. Boucsein. *Electrodermal activity*. Plenum Press, New York, 1992.

[5] J. Cacioppo and L. Tassinary. Inferring psychological significance from physiological signals. *American Psychologist*, 45:16–28, 1990.

[6] J. Chung and G. Scott Vercoe. The affective remixer: Personalized music arranging. In *CHI 2006: Conference on Human Factors in Computing Systems*, pages 393–398, Montréal, Canada, April 2006.

[7] M. Csíkszentmihályi. *Flow: The Psychology of Optimal Experience*. Harper Collins, Sussex, UK, 1990.

[8] P. Ekman. *Expression and the nature of emotion*. Erlbaum, Hillsdale, NJ, 1984.

[9] S. H. Fairclough. Fundamentals of physiological computing. *Interacting with Computers*, 21(1–2):133–145, 2009.

[10] N. H. Frijda. *The Emotions*. Cambridge University Press, New York, 1986.

[11] W. Härdle. *Smoothing Techniques, with Implementations in S*. Springer, New York, 1991.

[12] J. Healey, R. W. Picard, and F. Dabek. A new affect-perceiving interface and its application to personalized music selection. In M. Turk, editor, *Proceedings of the 1998 Workshop on Perceptual User Interfaces PUI*, page [online], San Francisco, CA, USA, 1998.

[13] C. Heinz and B. Seeger. Cluster kernels: Resource-aware kernel density estimators over streaming data. *IEEE Transactions on Knowledge and Data Engineering*, 20:880–893, 2008.

[14] E. Hollnagel. Is affective computing an oxymoron? *International Journal of Human-Computer Studies*, 59:65–70, 2003.

[15] D. Keltner and J. J. Gross. Functional accounts of emotions. *Cognition and Emotion*, 13:467–480, 1999.

[16] J. Kim and E. André. Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 30(12):2067–2083, 2008.

[17] J. T. Larsen, G. G. Berntson, K. M. Poelhman, T. A. Ito, and J. T. Cacioppo. *The psychophysiology of emotion*. Guilford, New York, 2007.

[18] M. Liljedahl, C. Sjömark, and N. Lefford. Using music to promote physical well-being via computer-mediated interaction. In *MusicNetwork Open Workshop, 5:5*, 2005.

[19] M. M. Llabre, S. B. Spitzer, P. G. Saab, G. H. Ironson, and N. Schneiderman. The reliability and specificity of delta versus residualized change as a measure of cardiovascular reactivity to behavioral challenges. *Psychophysiology*, 28:701–711, 1991.

[20] U. M. Nater, E. Abbruzzese, M. Krebs, and U. Ehlert. Sex differences in emotional and psychophysiological responses to musical stimuli. *International Journal of Psychophysiology*, 62:300–308, 2006.

[21] A. C. H. North and J. David. Musical preferences during and after relaxation and exercise. *American Journal of Psychology*, 113:43–67, 2000.

[22] A. C. H. North and D. J. Hargreaves. Responses to music in aerobic exercise and yogic relaxation classes. *British Journal of Psychology*, 87:535–547, 1996.

[23] N. Oliver and L. Kregor-Stickles. PAPA: Physiology and purpose-aware automatic playlist generation. In *ISMIR 2006: 7th International Conference on Music Information Retrieval*, pages 250–253, Victoria, Canada, October 2006.

[24] C. L. Pelletier. The effect of music on decreasing arousal due to stress: A meta-analysis. *Journal of Music Therapy*, 41:192–214, 2004.

[25] C. Peter and A. Herbon. Emotion representation and physiology assignments in digital systems. *Interacting with Computers*, 18:139–170, 2006.

[26] R. W. Picard. *Affective Computing*. MIT Press, Cambridge, MA, 1997.

[27] P. J. Rentfrow and S. D. Gosling. The do re mi's of everyday life: The structure and personality correlates of music preference. *Journal of Personality and Social Psychology*, 84:1236–1256, 2003.

[28] N. S. Rickard. Intense emotional responses to music: A test of the physiological arousal hypothesis. *Psychology of Music*, 32:371–388, 2004.

[29] D. A. Ritossa and N. S. Rickard. The relative utility of 'pleasantness' and 'liking' dimensions in predicting the emotions expressed by music. *Psychology of Music*, 32:5–22, 2004.

[30] J. A. Russell. Core affect and the psychological construction of emotion. *Psychological Review*, 110:145–172, 2003.

[31] A. Sanz and F. Villemarín. The role of perceived control in physiological reactivity: self-efficacy and incentive value as regulators of cardiovascular adjustment. *Biological Psychology*, 56:219–246, 2001.

[32] E. Schubert. The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music*, 35:499–515, 2007.

[33] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.

[34] J. A. Sloboda. *Exploring the musical mind: cognition, emotion, ability, function*. Oxford University Press, New York, 2005.

[35] E. M. Sokhadze. Effects of music on the recovery of autonomic and electrocortical activity after stress induced by aversive visual stimuli. *Applied Psychophysiology and Biofeedback*, 32:31–50, 2007.

[36] R. M. Stern, W. J. Ray, and K. S. Quigley. *Psychophysiological recording*. Oxford University Press, New York, 2001.

[37] A. Storr. *Music and the mind*. Free Press, New York, 1992.

[38] R. E. Thayer. *The Origin of Everyday Moods*. Oxford University Press, New York, 1996.

[39] N. Tractinsky. Tools over solutions? comments on interacting with computers special issue on affective computing. *Interacting with Computers*, 16(4):751–757, 2004.

[40] B. A. Turlach. Bandwidth selection in kernel density estimation: A review, 1993. Discussion Paper 9317, Institut de Statistique, Voie du Roman Pays 34, B-1348 Louvain-la-Neuve.

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

[42] G. D. Webster and C. G. Weir. Emotional responses to music: Interactive effects of mode, texture, and tempo. *Motivation and Emotion*, 29:19–39, 2005.

[43] R. Westermann, K. Spies, and G. Stahl. Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, 26:557–580, 1996.

[44] J. Wilder. *Stimulus and response: The law of initial values*. Wright, Bristol, UK, 1967.