AutoSelect: What You Want Is What You Get: Real-Time Processing of Visual Attention and Affect

Nikolaus Bee[1], Helmut Prendinger[2], Arturo Nakasone[3],
Elisabeth André[1], and Mitsuru Ishizuka[3]

[1] Institute of Computer Science, University of Augsburg
Eichleitnerstr. 30, D-86135 Augsburg, Germany
`nikolaus.bee@gmail.com, andre@informatik.uni-augsburg.de`
[2] National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
`helmut@nii.ac.jp`
[3] Graduate School of Information Science and Technology, University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
`arturo@mi.ci.i.u-tokyo.ac.jp, ishizuka@i.u-tokyo.ac.jp`

Abstract. While objects of our focus of attention ("where we are looking at") and accompanying affective responses to those objects is part of our daily experience, little research exists on investigating the relation between attention and positive affective evaluation. The purpose of our research is to process users' emotion and attention in real-time, with the goal of designing systems that may recognize a user's affective response to a particular visually presented stimulus in the presence of other stimuli, and respond accordingly. In this paper, we introduce the *AutoSelect* system that automatically detects a user's preference based on eye movement data and physiological signals in a two-alternative forced choice task. In an exploratory study involving the selection of neckties, the system could correctly classify subjects' choice of in 81%. In this instance of AutoSelect, the gaze 'cascade effect' played a dominant role, whereas pupil size could not be shown as a reliable predictor of preference.

1 Introduction and Motivation

A significant amount of recent research in human–computer interaction and interactive systems is driven by the aim of making the interaction with computers more intuitive, natural, and sensitive to the user's affective state. A prominent example is the multi-modal mixed-initiative dialogue system developed in the SmartKom project, where speech, gesture, and facial expression were synchronized, integrated, and interpreted in real-time, yielding robust understanding of the user's information state and intention [Wahlster, 2003]. The core purpose of the SmartKom dialogue system was to support users in task-oriented and collaborative applications, e.g. for receiving directions at an information kiosk.

A key assumption of interactive systems like SmartKom is that users have a particular (communication) goal in mind, which they want to achieve by successful interaction with the system, primarily through verbal utterances. Our

work relaxes this assumption in that we do not assume that users express their cognitive state verbally. Instead, we will rely solely on a user's unconscious (or non-conscious) and non-cognitive behavior, specifically eye movements and autonomic nervous system activity, which provides rich evidence on the user's focus (and shift) of visual attention [Duchowski, 2003] and affective state [Levenson, 1988], respectively.

We believe that visual attention and physiology based interactive technology is of high relevance to infotainment and e-learning applications, handicapped people, as well as to future interfaces as propagated in various ambient intelligence and ubiquitous computing projects [Streitz and Nixon, 2005]. In fact, many decisions of our daily life cannot be easily explained in terms of overt reasoning on premises. In a restaurant, for instance, we choose between different types of dishes. Unless price or dietary considerations are of primary importance, our decision for a particular dish might be based on our taste, our expectation of a specific (eating) experience, or even our current mood. Similarly, we might not be able to explain why we choose a particular piece of clothes to wear. The same situation seems to hold true even for products that provide clearly distinguishable functionality, such as digital cameras or cellular phones. In this context, the Japanese term "Kansei" is often used, which refers to a person's feeling or image of a product [terremoto.net, 2006]. The related discipline of Kansei engineering, sometimes translated as "sensory engineering" or "emotional usability", involves the design of products that use sensory attributes in order to elicit desired consumer responses. We take the experience of a product's Kansei as further evidence for the importance of interactive technologies that are aware of a person's affective evaluation of a product, and might respond accordingly, e.g. by providing further information on the product or by reinforcing a customer's non-conscious (positive) evaluation.

In this paper, we will describe a model (and system) that aims at detecting a user's positive affective evaluation of a visually presented item (preference) based on eye movements and physiological signals. Our notion of "positive affective evaluation" is loosely related to the notions of "liking" and "wanting", but also "interest". However, given that those notions have several meanings in the psychological literature [Feldman-Barrett, 2006], we preferred to use a more neutral term that is intended to refer to non-conscious biasing steps involved in decision making [Bechara et al., 1997]. The key technical challenge of our work is to process and interpret physiological signals and eye movements in real-time, given their individual characteristics, such as different latency and duration. Although affective computing [Picard, 1997] is an active research area, quite surprisingly, approaches to fusing physiological signals in real-time are almost non-existent (see [Prendinger and Ishizuka, 2005] for an early attempt).

In the rest of the paper, we first describe relevant background research, and then propose our own model for affective evaluation that is based on biometric signals and eye movements as input modalities. We will demonstrate the operation of our model by implementing an automatic necktie selector, and present our initial results that were obtained from an exploratory study.

2   Background and Related Work

In this section, we will report on related work that aims at interpreting physiological information as affective states, and eye movements as an indicator for preference. Recently, many studies were conducted to recognize affective states, typically the six basic emotions proposed by [Ekman, 1992] (or a subset of them), based on speech, facial expression, and physiological signals, or a combination of signals [Busso et al., 2004]. Others used physiological signals, sometimes in combination with speech, to detect emotions based on the two-dimensional model of [Lang, 1995] (see [Kim et al., 2005]). It is important to note that most of those approaches are *offline* methods that use feature extraction and machine learning techniques for affect recognition. We wish to emphasize that our goal is different in that we want to recognize affective states, specifically (positive) affective evaluation, in *real-time*, and therefore will rely on a knowledge-based approach that is informed by the psychophysiological literature.

2.1   Biometric Signals and Affective Evaluation

There are two famous emotion models available for real-time affect recognition from physiological signals, the *two-dimensional emotion* model [Lang, 1995] and the *autonomic specificity of emotions* approach [Levenson, 1988].

The two-dimensional emotion model advocated by [Lang, 1995] claims that all emotions can be characterized by two bipolar, but independent dimensions: (i) (judged) valence (pleasant or unpleasant, or: positive or negative); (ii) arousal (calm or aroused). Here, named emotions can be conceived as coordinate points in the arousal–valence space. While this model works well for basic emotions of Ekman, such as happy or sad, it is not obvious how to determine the location of positive affective evaluation (or other affect-related states such as confusion).

[Levenson, 1988], on the other hand, argues that (some) emotions can be distinguished by their associated pattern of autonomic nervous system activity, i.e. (some) emotions have "autonomic signatures". For instance, in an early study, [Ekman et al., 1983] found that there was a larger increase of heart rate with anger and fear than with happiness, and that skin temperature decrease was stronger with anger than with happiness, among other findings.

Unfortunately, literature on directly relating biometric signals to affective evaluation is sparse. The following paragraphs summarize findings from the literature. (Unless otherwise indicated, results are based on [Andreassi, 2000]).

Galvanic Skin Response (GSR). The GSR signal is an indicator of skin conductance (SC), which increases linearly with a person's level of overall arousal or stress. While an increased level of SC is associated with an orienting response to new or interesting stimuli, a very high level of SC is a good indicator of (negative) arousal (stress). [Bechara et al., 1997] also report on an interesting relation between SC responses and conscious decision making. In their study using a gambling task, (non-patient) subjects showed *anticipatory* SC responses before they knew explicitly that their choice was risky (as opposed to the patient subjects with prefrontal damage).

Blood Volume Pulse (BVP). The BVP signal is an indicator of blood flow. Since each heart beat (or pulse) presses blood through the vessels, BVP can also be used to calculate heart rate and inter-beat intervals. Heart rate increases with negatively valenced emotions, such as anxiety or fear.

Pupillary Response. Pupil size is affected by human emotional and cognitive processes [Hess, 1972]. Increase in pupil size is a good indicator of novelty, interest, and positive evaluation, but also of cognitive load, whereas decrease in pupil size indicates increased fatigue, and possibly negative stimuli ("perceptual avoidance"). However, [Hess, 1972, p. 509] also reports on subjects whose pupils dilated when shown pictures containing 'shock' content. He further speculates that increased levels of skin conductance might have an impact on pupil dilation. Presumably the most important finding for our research is the significant correlation between persons' (actual) attitude towards items such as food and consumer goods and pupil dilation. [Krugman, 1964] conducted studies demonstrating that goods (e.g. silverware patterns and greeting cards) inducing pupil responses indeed outsold goods without such responses.

Eye Blinks (EB). EBs occur throughout the day, with an average of 15–20 times per minute for a relaxed person. From a physiological point of view, only 2–4 blinks are necessary for an adult; while reading, the blink rate can drop to three blinks per minute. From a psychological perspective, blink frequency reflects negative affective states, such as nervousness, stress, and fatigue. Eye blink magnitudes were shown to be larger and latencies faster during negative as opposed to positive imagery. Moreover, higher arousal resulted in larger magnitude and shorter latency of eye blinks.

For detecting affective evaluation of visually presented stimuli, eye movements are of particular importance, and will be discussed next.

2.2   Eye Movements and Preference Formation

When presenting pairs of human faces to subjects and giving the instruction to decide on their attractiveness, [Shimojo et al., 2003] observed a phenomenon they called gaze 'cascade effect'. This phenomenon involves the gradual gaze shift toward the face that was eventually chosen (as more attractive), while gaze bias was initially distributed evenly between the two presented faces. The results of the two-alternative forced choice (2AFC) task used in their study demonstrated a progressive bias in subjects' gaze toward the chosen stimulus (preference formation), which was measured by the gaze time spent on the selected stimulus. However, the strong correlation between choice and gaze duration occurred only in the last one and half seconds before the decision was made.

A finding that [Shimojo et al., 2003] declared as surprising relates to the result that a larger cascade effect was found in the 'difficult' task, where the comparison between the attractiveness of faces was difficult, while intuitively, subjects were expected to more evenly distribute their gaze between stimuli in this case, in order to compare stimuli in as much detail as possible. The result was be explained by a theory claiming that gaze would significantly contribute

to decision-making when cognitive bias is weak. The importance of this result for our research derives from the fact that a large number daily choices, e.g. regarding consumer products, are also deficient of a strong cognitive bias, and hence contributes to the importance of investigating non-conscious human decisions.

We recently proposed an approach to estimating user interest that is based on both physiological signals and eye movements [Kon et al., 2006]. In order to achieve higher accuracy of estimating interest, gaze duration time was combined with skin potential level (SPL) based arousal detection. Here, arousal data allowed for more precise segregation into interest and non-interest regions in a two-dimensional space by using a machine learning method. A shortcoming of this approach was that SPL data were not aligned with gaze direction, such that we could not determine the visual 'source' of SPL changes. We will address this problem by explaining our model in the following section.

3   A Model for Affective Evaluation

The proposed model is, to our knowledge, the first attempt to estimate (positive) affective evaluation (preference) from physiological information (skin conductance, blood volume pulse, pupil size, blink rate) and focus of attention. The current model is confined to situations where the user is shown two visual stimuli (pictures), one to the left and one to the right of the screen, from which one will be chosen automatically depending on the user's gaze and physiological behavior. This can be conceived as an automated version of the two-alternative forced choice task described in [Shimojo et al., 2003]. It should be noted that most previous research on emotion recognition used stimuli with a single fixed location (showing e.g. pictures or videos) to elicit affective responses.

A key problem is to attribute affective meaning to visual stimuli, i.e. modeling a user's affective evaluation, given an ongoing stream of physiological activity. Typically, a person who chooses between two pictures will alternately look at the stimuli while physiological changes might occur. In order to explain the notion of *temporal matching* between physiological activity and visual stimuli, we will introduce the notions of 'time window' and '(un)directed signal'.

A *time window* refers to a temporal segment of sufficient duration to estimate relevant changes in bio-signal activity. Time windows differ depending on the latency (duration of onset), apex, and offset of each signal. Pupil size can be considered as a *directed* signal, since it reflects physiological changes directly influenced by the stimulus attended to. On the other hand, SC, BVP, and EBs are seen as *undirected* signals in our setting (where persons will select from two stimuli); i.e. the time window for a significant rise of the level of SC may not be included in the time period a person is attending to one single stimulus, but may become evident after multiple shifts between the two stimuli. Our model for affective preference is encoded as a Bayesian network using Netica$^{\text{TM}}$ software from [Norsys, 2003] (see Fig. 1).

The network is based on the findings reported in Sects. 2.1 and 2.2. Specifically, the "Gaze Bias" node encodes the gaze cascade effect that was determined
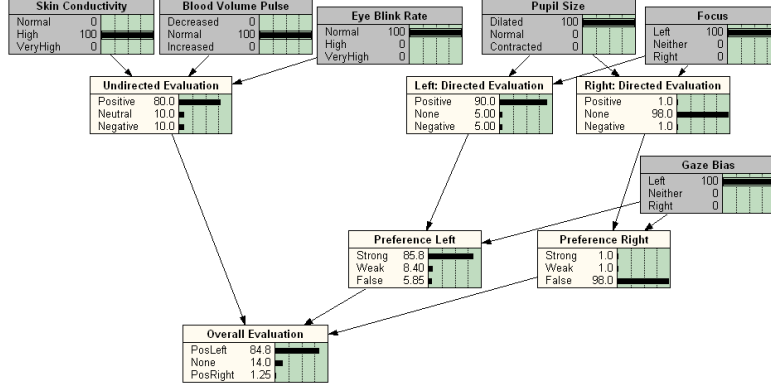
Fig. 1. Bayesian network for recognizing affect based preference for the two stimuli case, assumed as shown to the left and to the right of a screen

empirically in [Shimojo et al., 2003], i.e. the value is "Left" ("Right") if the user continuously fixates the left (right) stimulus for more than one second. A threshold rules out cases where users initially fixate a stimulus longer without having inspected the other stimulus.

Undirected evaluation may contribute to the overall evaluation (preference). If positive, this is seen as evidence that the selected stimulus received (positive) approval. A negative value is assumed to contribute to the (negative) evaluation of the non-selected stimulus. Those interpretations are model assumptions and highly speculative. It is important to notice that the instance of the AutoSelect system (as a tie selector) depends primarily on gaze. However, affect-related signals were included in the model for the more general case of automatically selecting between stimuli that are more emotionally arousing than ties.

4   Algorithms for Real-Time Processing of Affect and Visual Attention

Online feature analysis requires methods that are different from those commonly used in off-line processing since we cannot rely on the availability of the entire signal history built up during some interaction session, including (global) baselines, minima, or maxima.

A core technique underlying most of our algorithms is the derivative as we are only interested in changes over time. We consider different time windows for detecting tendencies (BVP rate, 5 s; eye blink rate, 10 s; pupil response, 1 s) and counting absolute values (skin conductance startle, 1 s; focus of attention, 300 ms). Our analysis is thus independent from a global baseline taken prior to the interaction session. For the task of selecting ties, it was not deemed necessary

to employ a dynamic Bayesian network, and hence each choice is treated as independent from the previous one.

Our algorithms partly depend on the hardware available to us. We used the ProComp Infinity^{TM} encoder from [Thought Technology, 2005] to process bio-signals, and faceLAB^{TM} v4 from [Seeing Machines, 2005] for eye related processing. The ProComp Infinity^{TM} encoder was set to 20 samples/s that was sufficiently high for the investigated signals, e.g. pupil change occurs with 200 ms latency and develops a local maximum within 1500 ms. The faceLAB^{TM} eye tracker has a sampling rate of 60 Hz. Calculation of the head and eye positions with the faceLAB^{TM} software amounted to an average latency of 30 ms.

Skin Conductance. For arousal detection, we modified the algorithm originally developed by [Healey, 2000] by removing slight jitters through smoothing the signal with a basic mean filter. For robust threshold-based startle detection, the square of the raw smoothed signal was calculated. A further improvement was achieved by adapting the threshold based on the past signal during operation.

Blood Volume Pulse. As BVP depends on time and is subject to rise and fall, the derivative of the signal was used as an index of increased heart rate. As soon as the derivative exceeds a certain threshold, a heart beat is detected. Robustness was increased by using a heuristic that is based on the assumption that there is no pulse rate below 45 or above 120 beats/minute.

Eye Blink Rate. Eye blink frequency is obtained from eye closure which is calculated by the faceLAB^{TM} software. If the eye is closed for less than one second a blink is detected. Eye blinks last for 400 ms on average.

Pupil Size. The faceLAB^{TM} software also provides information about the pupil size. Using the method proposed in [Schultheis and Jameson, 2004], artifacts due to EBs are eliminated by suppressing the signal 200 ms before and 1 s after a blink. Since our interest is in the dilation and contraction of the pupil size over time, the derivative is used as a feature.

Gaze Bias and Focus. The gaze bias is calculated from the gaze point distribution regarding two visually presented objects. [Shimojo et al., 2003] analyzed the gaze cascade effect by considering the last 1.5 seconds before a decision was made. We adopted this value for our calculation of gaze bias. Focus was calculated by counting gaze points in time windows of 300 ms. (The implementation of a fixation algorithm was not mandatory for our application.)

5   Exploratory Study

A system that may automatically detect users' choices seems to break new ground. We therefore conducted an exploratory study using the AutoSelect system. Our first application is an automatic necktie selector, where subjects are shown a pair of ties and the AutoSelect system tries to detect the preferred tie. Subjects were given no instruction other than having to choose a tie for themselves or their friend for a graduation party.

5.1   Procedure and Design

Eight subjects (4 female, 4 male), all students or researchers from NII, participated and received an award of 1,000 Yen. Subjects entered the experimental room individually and were provided written instructions about their task.

Subjects were seated in front of an 20.1 inch display with attached infrared lights (see Fig. 2) and their head and eyes were calibrated. This procedure has to be performed for each individual once, and takes approximately 5 minutes. (If a person wears glasses, the procedure might be prolonged to due reflection from the infrared light.) Next, the biometrical sensors of the ProComp Infinity encoder were attached to the subject. The SC sensor was attached to the index finger and the ring finger of the non-dominant hand, and the BVP sensor was fixed to the thumb of the same hand.

Fig. 2. Experimental setup

A session was initialized by subjects pressing a 'start' button in a web page based interface. Then the following procedure was iterated for 62 pairs of ties. First, a center located 'dot' was shown on the screen for 2.5 s in order to eliminate any initial gaze bias. Next, a pair of ties was presented, located left and right on the screen. In order to guarantee that subjects actually compare the ties, automatic selection was suppressed within the first 2.5 s. (This value was based on the empirically determined decision time of 4 s in [Shimojo et al., 2003].) After the system decision, the selected tie was presented and subjects were asked to indicate whether the system choice is correct by clicking on a 'yes' or 'no' button. Then the next iteration started with the initial view of a center dot.

One initial set of 32 tie pairs was prepared, and the chosen ties were put back into the tie pool, which was used to create the subsequent set of 16 pairs, and so on. Eventually, subjects were shown a single pair of ties they presumably liked best. Hence, subjects were exposed to 63 pairs and performed 62 decisions in total. In the initial set of tie pairs, two partitions were created with 13 pairs each. One partition contained pairs of 'different' type ties, i.e. formal (decent) vs. 'entertainment' (adventurous) style ties, whereas the other partition contained 'similar' type ties that differed only in color or had a slightly different pattern but the same color. The motivation of this grouping was to investigate differences in subjects' decision behavior for presumably 'easy' vs. 'hard' decisions. All sessions were logged and lasted for about 10 minutes.

5.2   Results

The primary result concerns the classification accuracy of the AutoSelect system. In our study, the system could detect subjects' choices correctly in 81% of the cases. The worst recognition rate was 68%. Given a chance level of 50%, the

system performed very well. (One subject was excluded from analysis because of distorted values due to starting a conversation during the experiment.) It is noted that we did not aim at validating the network in which case we would simply have compared user decisions with system decisions. Instead, we wanted to investigate the users' interactive experience with a running system, which can reveal e.g. issues related to the latency between user decision and system decision. Informal comments on the system indeed indicated that subjects were surprised about the system's reliability to timely identify which tie they liked more. Some of the misclassifications were related to a design problem, i.e. when subjects moved their face out of the camera range. The next version of AutoSelect should alert subjects in those situations.

In the following, we will discuss results related to gaze and bio-signals. As to gaze, we were particularly interested in results comparable to the 'difficult' vs. 'easy' choice finding reported in [Shimojo et al., 2003]. We hence compared recognition rates and decision times for 'different' vs. 'similar' tie pairs (see Fig. 3). In line with Shimojo, Simion and colleagues, the decision time for different ties was significantly longer than for similar ties ($t(180) = -1.66$; $p = 0.049$). We used a one-tailed t-test assuming unequal variances in our analysis. This result supports the hypothesis that a choice between unlike items relies on (time consuming) cognitive processing, whereas similar items might be chosen based on non-conscious ('intuitive') preference. We also note that the system calculated the choice between similar ties more accurately.

Although of secondary importance (and certainly speculative given the small number of subjects), we also investigated gender differences in terms of recognition rate and decision time (see Fig. 4). As opposed to a sometimes heard (joking) prejudice, female subjects seemingly decided significantly faster than male subjects ($t(432) = -3.79$; $p = 0.00009$). We should emphasize that this result relates to the decision time calculated by the system, and not necessarily to the decision time of the subjects. It can also be observed that the recognition rate is higher for female subjects.

Our analysis of biometric activity in this paper is confined to the directed signal of pupil size, which is closely related to perception but also to emotion (valence). Based on the findings in [Krugman, 1964, Hess, 1972], we predicted that pupil dilation would occur whenever subjects view the tie they declared as their choice. (The data of one subject were accidentally overwritten and hence missing.) However, only three subjects (S4, S6, S7) supported this hypothesis,

	rec. rate / dec. time (s)	
different	75%	6.80
similar	81%	7.65
overall	**78%**	**7.22**

	rec. rate / dec. time (s)	
female	89%	6.15
male	75%	7.43
overall	**81%**	**6.88**

Fig. 3. Recognition results and mean decision time for similar/different type ties

Fig. 4. Recognition results and mean decision time dependent on gender

Subj	C	non-C	t(df)	df	p
S3	3.08	3.13	−3.50	4612	0.0002
S4	3.23	3.19	2.43	5098	0.0076
S5	2.35	2.32	0.91	5680	0.179
S6	3.54	3.51	2.41	4114	0.008
S7	3.27	3.22	5.50	7358	<0.0001
S8	3.13	3.14	−0.69	5393	0.75

Subj	C	non-C	t(df)	df	p
S3	3.02	3.10	−3.60	2443	<0.0001
S4	3.17	3.24	−3.49	1785	0.0003
S5	2.19	2.30	−2.38	3378	0.0087
S6	3.55	3.56	−1.01	2829	0.155
S7	3.34	3.24	7.83	3414	<0.0001
S8	3.13	3.18	−1.03	1737	0.15

Fig. 5. One-tailed t-test for pupil size of preferred (choice) and non-choice ties (in millimeter)

Fig. 6. One-tailed t-test for pupil size of choice vs. non-choice ties when seen for the first time (in millimeter)

whereas S3 supports the opposite, i.e. a significant contraction of pupil size for the preferred tie, while the results for S5 and S8 are not significant (see Fig. 5). Total dwell time on the preferred tie was always greater than on the non-choice tie, with a significant difference in four cases. This result, however, is not surprising given that the gaze cascade effect implies that more time will eventually be spent on the choice.

In view of those results, our initial guess was that *habituation*, i.e. the progressive diminution of behavioral response probability with repetition of a stimulus, might have been responsible for the non-significance of some data. We hence analyzed only those situations where the subjects saw the ties for the first time (ties in the initial 32 pairs). Much to our surprise, we found the opposite effect prevailing, i.e. pupil size for the preferred tie was mostly smaller than for the choice tie, and in three cases (S3, S4, S5), significantly smaller (see Fig. 6).

Finally, we tested a conjecture put forth in [Simion, 2005], namely that pupil size might be a physiological precursor of preference or choice *before* any gaze particular pattern occurs. Specifically, pupil size and dwell time was compared for choice and non-choice ties for the single instance where subjects looked at a tie for the (very) first time. (Note that the analysis in the previous paragraph included multiple situations effected by the comparative task.) However, only two subjects showed significantly smaller pupil size for the preferred tie, and one subject had significantly longer dwell time for the choice.

The present study could thus not establish pupil size (or duration) as a reliable predictor for preference, and leaves open many questions. Looking at the pupil size literature [Hess, 1972], it can be observed that studies often used 'strong' visual stimuli, including the display of 'shock content' but also nudity. The stimuli chosen for our study (neckties) are obviously of a different quality. (URL: `http://research.nii.ac.jp/~prendinger/autoselect`)

6   Conclusions

The contribution of this paper is two-fold. First, we introduced a model for the joint interpretation of attention and affect. The model is encoded as a Bayesian

network and may infer a user's affective evaluation or preference based on eye movements and bio-signal activity. The model is used for AutoSelect, a system that primarily exploits the gaze cascade effect in order to predict the choice of users in a two-alternative forced choice task. Biometric signals may change the probability of the positive left (right) choice in the network, but not the computed preference for the 'left' or 'right' stimulus.

Second, we conducted an exploratory study to test whether AutoSelect can correctly predict the choice of a user. The accuracy of the system with a limited number of subjects (7) is reasonably high (81%). Pupil size could not be established as a reliable predictor of affective evaluation in our study. Certainly, if our sole interest would be tie selection, biometric signals could be removed from the network altogether. However, we plan a follow-up study that employs life-like characters as interaction partners [Prendinger and Ishizuka, 2004] and thus provides an emotionally more arousing interaction than the tie selection task. In this way, we hope to obtain a clearer picture of the utility of pupil size and other biometric input signals, which were of little importance in the tie selection task.

Acknowledgments

References

[Andreassi, 2000] Andreassi, J. L. (2000). *Psychophysiology. Human Behavior & Physiological Response.* Lawrence Erlbaum Associates, Mahwah, NJ, 4 edition.

[Bechara et al., 1997] Bechara, A., Damasio, H., Tranel, D., and Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275:1293–1295.

[Busso et al., 2004] Busso, C., Deng, Z., Yildirim, S., Buut, M., Lee, C. M., Kazemzadeh, A., Lee, S., Neumann, U., and Narayanan, S. (2004). Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proceedings of 6th International Conference on Multimodal Interfaces (ICMI-04)*, pages 205–211. ACM Press, New York.

[Duchowski, 2003] Duchowski, A. T. (2003). *Eye Tracking Methodology: Theory and Practice.* Springer, London, UK.

[Ekman, 1992] Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4):169–200.

[Ekman et al., 1983] Ekman, P., Levenson, R. W., and Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221:1208–1210.

[Feldman-Barrett, 2006] Feldman-Barrett, L. (2006). Emotions as natural kinds? *Perspectives on Psychological Science*, 1:28–58.

[Healey, 2000] Healey, J. A. (2000). *Wearable and Automotive Systems for Affect Recognition from Physiology*. PhD thesis, Massachusetts Institute of Technology.

[Hess, 1972] Hess, E. H. (1972). Pupillometrics: A method of studying mental, emotional and sensory processes. In Greenfield, N. and Sternbach, R., editors, *Handbook of Psychophysiology*, pages 491–531. Holt, Rinehart & Winston, New York.

[Kim et al., 2005] 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 9th European Conference on Speech Communication and Technology*.

[Kon et al., 2006] Kon, M., Koshizen, T., and Prendinger, H. (2006). A new user–machine interface using cross-modal computation for deep interest estimation. Towards quantifying user satisfaction. In *Proceedings of IUI-06 Workshop on Effective Multimodal Dialogue Interfaces*, pages 25–34.

[Krugman, 1964] Krugman, H. (1964). Some applications of pupil measurement. *Journal of Marketing Research*, 1:15–19.

[Lang, 1995] Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5):372–385.

[Levenson, 1988] Levenson, R. W. (1988). Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity. In Wagner, H. L., editor, *Social Psychophysiology and Emotion: Theory and Clinical Applications*, pages 17–42. John Wiley & Sons, Hoboken, NJ.

[Norsys, 2003] Norsys (2003). Norsys Software Corp. Netica.
URL: http://www.norsys.com.

[Picard, 1997] Picard, R. W. (1997). *Affective Computing*. The MIT Press, Cambridge, MA.

[Prendinger and Ishizuka, 2004] Prendinger, H. and Ishizuka, M., editors (2004). *Life-Like Characters. Tools, Affective Functions, and Applications*. Cognitive Technologies. Springer Verlag, Berlin Heidelberg.

[Prendinger and Ishizuka, 2005] Prendinger, H. and Ishizuka, M. (2005). The Empathic Companion: A character-based interface that addresses users' affective states. *International Journal of Applied Artificial Intelligence*, 19(3):267–285.

[Schultheis and Jameson, 2004] Schultheis, H. and Jameson, A. (2004). Assessing cognitive load in adaptive hypermedia systems: Physiological and behavioral methods. In *Proceedings Adaptive Hypermedia and Adaptive Web-based Systems (AH-04)*, pages 225–234, Berlin. Springer.

[Seeing Machines, 2005] Seeing Machines (2005). Seeing Machines.
URL: http://www.seeingmachines.com/.

[Shimojo et al., 2003] Shimojo, S., Simion, C., Shimojo, E., and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, 6(12):1317–1322.

[Simion, 2005] Simion, C. (2005). *Orienting and Preference: An Enquiry into the Mechanisms Underlying Emotional Decision Making*. PhD thesis, California Institute of Technology.

[Streitz and Nixon, 2005] Streitz, N. and Nixon, P. (2005). The Disappearing Computer. Guest editors' introduction to Special Issue. *Communications of the ACM*, 48:33–35.

[terremoto.net, 2006] terremoto.net (2006). Kansei engineering: Incorporating affection and emotion into the design process. URL: `http://terremoto.net/kansei/`.

[Thought Technology, 2005] Thought Technology (2005). Thought Technology Ltd. URL: `http://www.thoughttechnology.com`.

[Wahlster, 2003] Wahlster, W. (2003). Towards symmetric multimodality: Fusion and fission of speech, gesture and facial expression. In *Proceedings 26th German Conference on Artificial Intelligence*, pages 1–18, Berlin Heidelberg. Springer LNAI 2821.