Automatic Measurement of Affect in Dimensional and Continuous Spaces: Why, What, and How?*

Hatice Gunes
Imperial College London, U.K.
University of Technology Sydney, Australia
h.gunes@imperial.ac.uk

Maja Pantic
Imperial College London, U.K.
University of Twente, The Netherlands
m.pantic@imperial.ac.uk

ABSTRACT
This paper aims to give a brief overview of the current state-of-the-art in automatic measurement of affect signals in dimensional and continuous spaces (a continuous scale from -1 to +1) by seeking answers to the following questions: i) *why* has the field shifted towards dimensional and continuous interpretations of affective displays recorded in real-world settings? ii) *what* are the affect dimensions used, and the affect signals measured? and iii) *how* has the current automatic measurement technology been developed, and *how* can we advance the field?

Author Keywords
Automatic measurement of human affect, dimensional and continuous affect recognition, multicue and multimodal affect recognition.

ACM Classification Keywords
A.1 Introduction and survey, H.1.2 User/machine systems: Human information processing, I.5.4 Pattern recognition applications

WHY MEASURE AFFECT IN DIMENSIONAL SPACES?

According to research in psychology, three major approaches to affect modelling can be distinguished [10]: categorical, dimensional, and appraisal-based approach. The categorical approach claims that there exist a small number of emotions that are basic, hard-wired in our brain, and recognized universally (e.g. [5]). This theory on universality and interpretation of affective nonverbal expressions in terms of basic emotion categories has been the most commonly adopted approach in research on automatic measurement of human affect.

However, a number of researchers have shown that in everyday interactions people exhibit non-basic, subtle and rather complex affective states like thinking, embarrassment or depression. Such subtle and complex affective states can be expressed via dozens of anatomically possible facial and bodily expressions, audio or physiological signals. Therefore, a single label (or any small number of discrete classes) may not reflect the complexity of the affective state conveyed by such rich sources of information [23]. Hence, a number of researchers advocate the use of dimensional description of human affect, where affective states are not independent from one another; rather, they are related to one another in a systematic manner (e.g., [10, 23, 24]).

It is not surprising, therefore, that automatic affect sensing and recognition researchers have recently started exploring how to model, analyse and interpret the subtlety, complexity and continuity (represented along a continuum from -1 to +1, without discretisation) of affective behaviour in terms of latent dimensions, rather than in terms of a small number of discrete emotion categories.

The most widely used dimensional model is Russell's two-dimension 'circumplex model of affect', where emotions are seen as combinations of arousal and valence [23].

Scherer and colleagues introduced another set of psychological models, referred to as componential models of emotion, which are based on appraisal theory [7, 10, 24]. In the appraisal-based approach emotions are generated through continuous, recursive subjective evaluation of both our own internal state and the state of the outside world (relevant concerns/needs) [7, 8, 10, 24]. How to use the appraisal-based approach for automatic measurement of affect is an open research question as this approach requires complex, multicomponential and sophisticated measurements of change. One possibility is to reduce the appraisal models to dimensional models (e.g., 2D space of arousal-valence).

Another model, known as OCC [19] is also established as a standard cognitive appraisal model for emotions, and has mostly been used in affect synthesis (in embodied conversational agent design).

---

*An extended version of this paper has been published in the International Journal of Synthetic Emotions, in Jan. 2010.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. For any other use, please contact the Measuring Behavior secretariat: info@measuringbehavior.org.

WHAT ARE THE AFFECT DIMENSIONS AND SIGNALS USED FOR AUTOMATIC MEASUREMENT?

An individual's inner emotional state may become apparent by subjective experiences (how the person feels), external/outward expressions (audio/visual signals), and internal/inward expressions (bio signals). However, these may be incongruent, depending on the context (e.g., feeling angry and not expressing it outwardly). This poses a true challenge to automatic sensing and analysis.

Currently, a number of affect recognisers attempt to label both the felt (e.g., [1, 17]) and the internally/externally expressed (e.g., [14, 15]) emotions.

Affect Dimensions

Despite the existence of the abovementioned emotion models, in automatic measurement of dimensional and continuous affect, *valence* (how positive or negative the affect is), *activation* (how excited or apathetic the affect is), *power* (the sense of control over the affect), and *expectation* (the degree of anticipating or being taken unaware) appear to make up the four most important affect dimensions [7]. Although ideally the intensity dimension could be derived from the other dimensions, to guarantee a complete description of affective colouring, some researchers include *intensity* (how far a person is away from a state of pure, cool rationality) as the fifth dimension. However, search for optimal low-dimensional representation of affect remains open [7].

Visual Signals

Facial actions (e.g., pulling eyebrows up) and facial expressions (e.g., producing a smile), and to a much lesser extent bodily postures (e.g., backwards head bend and arms raised forwards and upwards) and expressions (e.g., head nod), form the widely known and used visual signals for automatic affect measurement. Dimensional models are considered important in this task as a single label may not reflect the complexity of the affective state conveyed by a facial expression, body posture or gesture. Ekman & Friesen [6] considered expressing discrete emotion categories via face, and communicating dimensions of affect via body as more plausible. A number of researchers have investigated how to map various visual signals onto emotion dimensions. For instance, [23] mapped the facial expressions to various positions on the two-dimensional plane of arousal-valence, while [4] investigated the emotional and communicative significance of head nods and shakes in terms of arousal and valence dimensions, together with dimensional representation of solidarity, antagonism and agreement.

Audio Signals

Audio signals convey affective information through explicit (linguistic) messages, and implicit (acoustic and prosodic) messages that reflect the way the words are spoken. There exist a number of works focusing on how to map audio expression to dimensional models. Cowie et al. used valence-activation space (similar to valence-arousal) to model and assess affect from speech [2, 3]. Scherer and colleagues have also proposed how to judge emotion effects on vocal expression, using the appraisal-based theory [10].

Bio Signals

The bio-signals used for automatic measurement of affect are galvanic skin response that increases linearly with a person's level of arousal [1], electromyography (frequency of muscle tension) that is correlated with negatively valenced emotions [13], heart rate that increases with negatively valenced emotions such as fear, heart rate variability that indicates a state of relaxation or mental stress, and respiration rate (how deep and fast the breath is) that becomes irregular with more aroused emotions like anger or fear [1, 13]. Measurements recorded over various parts of the brain including the amygdala also enable observation of the emotions felt [22]. For instance, approach or withdrawal response to a stimulus is known to be linked to the activation of the left or right frontal cortex, respectively. A number of studies also suggest that there exists a correlation between increased blood perfusion in the orbital muscles and stress levels for human beings. This periorbital perfusion can be quantified through the processing of thermal video (e.g., [26]).

HOW IS THE CURRENT TECHNOLOGY DEVELOPED?

Data Acquisition and Annotation

Cameras are used for acquisition of face and bodily expressions, microphones are used for recording audio signals, motion capture systems are utilized for recording 3D affective postures and gestures, and thermal (infrared) cameras are used for recording blood flow and changes in skin temperature. In the bio-signal research context, the subject being recorded usually wears a headband or a cap on which electrodes are mounted, a clip sensor, or touch type electrodes. The subject is then stimulated with emotionally-evocative images or sounds. Acquiring affect data without subjects' knowledge is strongly discouraged and the current trend is to record spontaneous data in more constrained conditions such as an interview setting, where subjects are still aware of placement of the sensors and their locations.

Annotation of the affect data is usually done separately for each modality assuming independency between the modalities. A major challenge is the fact that there is no coding scheme that is agreed upon and used by all researchers in the field that can accommodate all possible communicative cues and modalities. In general, the annotation tool Feeltrace is used for annotating the external expressions (audio and visual signals) with *continuous* traces (impressions) in dimensional spaces. Feeltrace allows observers to watch an audio-visual recording and move their cursor within the affect space to rate their impression about the affective state of the subject [2]. For annotating the internal expressions (bio signals), the level of valence and arousal is usually extracted from subjective experiences (subjects' own responses) (e.g., [17, 22]) due to the fact that feelings induced by an image can be very different from subject to subject. When discretised dimensional annotation is adopted (as opposed to continuous one), researchers seem to use different intensity levels: either a ten-point Likert scale (e.g., 0-low arousal, 9-high arousal) or a range between -1.0 and 1.0 (divided into a number of levels) [11]. The final annotation is usually calculated as the mean of the observers' ratings. Development of an easy to use, unambiguous and intuitive annotation scheme that is able to incorporate inter-observer agreement levels remains an important challenge in the field.

Obtaining high inter-observer agreement is another challenge in affect data annotation, especially when (continuous) dimensional approach is adopted. To date, researchers have mostly chosen to use self-assessments (subjective experiences, e.g. [13]) or the mean (within a predefined range of values) of the observers' ratings (e.g. [16]). Modelling inter-observer agreement levels within automatic affect analyzers, and finding which signals better correlate with self assessments and which ones better correlate with independent observer assessments remains as a challenging issues in the field.

Automatic Measurement of Affect in Continuous Spaces

After affect data has been acquired and annotated, representative and relevant features need to be extracted prior to the automatic measurement of affect in dimensional and continuous spaces. The feature extraction techniques used for each communicative source are similar to the previous works (reviewed in [12]) adopting a categorical approach to affect recognition.

There are a number of additional issues which need to be taken into account when applying a dimensional approach to affect recognition.

The interpretation accuracy of expressions and physiological responses in terms of continuous emotions is very challenging. While visual signals appear to be better for interpreting valence, audio signals seem to be better for interpreting arousal [11]. A thorough comparison between all modalities would indeed provide a better understanding of which emotion dimensions are better recognised from which modalities (or cues).

The window size to be used to achieve optimal affect recognition is another issue that the existing literature does not provide a unique answer to. Current affect recognizers employ various window sizes depending on the modality, e.g., 2-6 seconds for speech, 3-15 seconds for bio-signal [15]. There is no consensus on how the efficiency of such a choice should be evaluated.

Measuring the intensity of expressed emotion appears to be modality dependent. The way the intensity of an emotion is apparent from physiological data may be different than the way it is apparent from visual data. Moreover, little attention has been paid so far to whether there are definite boundaries along the affect continuum to distinguish between various levels or intensities. Currently intensity is measured by quantizing the affect dimensions into arbitrary number of levels such as neutral, low and high (e.g., [16, 17, 27]). Separate models are then built to discriminate between pairs of affective dimension levels, for instance, low vs. high, low vs. neutral, etc. Generalizing intensity analysis across different subjects is a challenge yet to be researched as different subjects express different levels of emotions in the same situation.

The Baseline problem is another major challenge in the field. For tactile modality (bio signals) this refers to the problem of finding a condition against which changes in measured physiological signals can be compared (a state of calmness). For audio modality this is usually achieved by segmenting the recordings into turns using energy based voice activity detection and processing each turn separately. For visual modality the aim is to find a frame in which the subject is expressionless and against which changes in subject's motion, pose, and appearance can be compared. This is achieved by manually segmenting the recordings, or by constraining the recordings to have the first frame containing a neutral expression. However, expecting expressionless state in each recording or manually segmenting recordings so that each segment contains a baseline expression are strong, unrealistic constrains for analysis and processing of affective information.

Feature space with high dimensionality hinders automatic affect measurement. For instance, various works in the field have reported that they extract 2,520 features for each frame of an input facial video, 4,843 features for each utterance, 16,704 EEG features for each stream etc. (see [11] for details). Having fewer training samples than features per sample impedes the learning of the target classification. Various dimensionality reduction or feature selection techniques have been applied (e.g., Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA), kernel PCA (KPCA) Sequential Backward Selection) to mitigate this problem. Creating dimensionality reduction techniques with specific applications to automatic measurement of affect in dimensional and continuous spaces remains as an issue to be explored.

Generalisation capability of automatic affect analysers across subjects is still a challenge in the field. Kulic & Croft [17] reported that for bio-signal based affect measurement subjects seem to vary not only in terms of response amplitude and duration, but for some modalities, a number of subjects show no response at all. This makes generalisation over unseen subjects a very difficult problem. When it comes to other modalities, most of the works in the field report only on subject dependent dimensional affect measurement and recognition due to limited number of subjects and data (e.g., [27]).

Modality fusion refers to combining and integrating all incoming unimodal events into a single representation of the affect expressed by the user. When it comes to integrating multiple modalities, the major issues are: i) when to integrate the modalities (at what abstraction level to do the fusion), ii) how to integrate the modalities (which criteria to use), iii) how to deal with the increased number of features due to fusion, iv) how to deal with the asynchrony between the modalities (e.g., video is recorded at 25 Hz, audio is recorded at 48 kHz while EEG is recorded at 256-512 Hz), and v) how to proceed with fusion when there is conflicting information conveyed by the modalities. Despite a number of efforts in the discrete affect recognition field (reviewed in [12]), these issues remain yet to be explored for dimensional and continuous affect recognition.

Classification methods used for dimensional and continuous affect measurement should be able to produce continuous values for the target dimensions. Some of the classification schemes that have been explored for this task are, namely, Support Vector Regression (SVR), Conditional Random Fields (CRF), and Long Short-Term Memory Recurrent Networks (LSTM-RNN). Overall, there is no agreement on how to model dimensional affect space (continuous vs. quantised) and which classifier is better suited for automatic, multimodal, continuous affect analysis using a dimensional representation. The design of emotion-specific classification schemes that can handle multimodal and spontaneous data is one of the most important issues in the field.

Evaluation measures applicable to categorical affect recognition are not directly applicable to dimensional approaches. Using the Mean Squared Error (MSE) between the predicted and the actual value of arousal and valence, instead of the recognition rate (i.e., percentage of correctly classified instances) is the most commonly used measure by related work in the literature (e.g., [14, 27]). However, using MSE might not be the best way to evaluate the performance of dimensional approaches to automatic affect measurement and recognition. Therefore, the correlation coefficient, that evaluates whether the model has managed to capture patterns inhibited in the data at hand, is also employed by several studies (e.g., [14, 18]) together with MSE. Overall, however, how to obtain optimal evaluation metrics for continuous and dimensional emotion recognition remains an open research issue [11].

HOW CAN WE ADVANCE THE FIELD?

The analysis provided in this paper indicates that the automatic affect sensing field has slowly started shifting from categorical (and discrete) affect recognition to dimensional (and continuous) affect recognition to be able to capture the complexity of affect expressed in *real* world settings, by the *real* people. Despite the existence of a number of dimensional emotion models, the two-dimensional model of arousal and valence appears to be the most widely used model in automatic measurement from audio, visual and bio signals. The current automatic measurement technology has already started dealing with spontaneous data obtained in less-controlled environments using various sensing devices, and exploring a number of machine learning techniques and evaluation measures. However, real-world settings pose many challenges to continuous affect sensing and recognition (e.g., when subjects are not restricted in terms of mobility, the level of noise in all recorded signals tends to increase).

To date, only a few systems have actually achieved dimensional affect recognition from multiple modalities. These are reviewed in [11]. Overall, existing systems use different training/testing datasets (which differ in the way affect is elicited and annotated), they differ in the underlying affect model (i.e., target affect categories) as well as in the employed modality or combination of modalities, and the applied evaluation method. As a consequence, it remains unclear which classification method is suitable for dimensional affect recognition from which modalities and cues. These challenges should be addressed in order to advance the field while identifying the importance and feasibility of the following issues. **1)** Among the available remotely observable and remotely unobservable modalities, which ones should be used for automatic dimensional affect recognition? Should we investigate the innate priority among the modalities to be preferred for each affect dimension? Does this depend on the context (who the subject is, where she is, what her current task is, and when the observed behaviour has been shown)? **2)** When labelling emotions, which signals better correlate with self assessment and which ones correlate with independent observer assessment? **3)** How does the baseline problem affect recognition? Is an objective basis (e.g., a frame with an expressionless display) strictly needed prior to computing the dimensional affect values? If so, how can this be obtained in a fully automatic manner from spontaneous data? **4)** How should intensity be modelled for dimensional and continuous affect recognition? Should the aim be personalizing systems for each subject, or creating systems that are expected to generalize across subjects? **5)** In a continuous affect space, how should duration of affect be defined? How can this be incorporated in automated systems? Will focusing on shorter or longer observations affect the accuracy of the measurement process?

Finding straightforward answers to these questions is beyond the scope of this paper. Although research fields such as engineering, computer science, psychology, neuroscience, and cognitive sciences seem to be somewhat detached and have their own research community and audience, emotion research is inherently multi-disciplinary. Great advances in emotion research are possible, however, depend on all the aforementioned fields stepping out of their labs, working side-by-side together in real-life applications, and sharing the experience and the insight acquired on the way, to make emotion research *tangible* for the *real world* and the *real people* [20]. Pioneering projects representing such inter-disciplinary effort have already started emerging, ranging, for instance, from publishing compilation books of related work papers (e.g., [9]) to projects as varied as affective human-embodied conversational agent interaction (e.g., European Union FP 7 SEMAINE [25]), and affect sensing for autism (e.g., [21]).

ACKNOWLEDGMENTS

The research of Hatice Gunes is funded by [FP7/2007-2013], grant agreement no 211486 (SEMAINE). The research of Maja Pantic is funded in part by the ERC Starting Grant agreement no. ERC-2007-StG-203143 (MAHNOB).

REFERENCES

1. Chanel, G., Ansari-Asl, K. & Pun, T. Valence-arousal evaluation using physiological signals in an emotion recall paradigm, *Proc. of IEEE SMC* (2007), 2662-2667.
2. Cowie, R., Douglas-Cowie, E., Savvidou, S., Mcmahon, E., Sawey, M. & Schröder, M. 'FEELTRACE': An instrument for recording perceived emotion in real time, *Proc. ISCA W. on Speech and Emotion* (2000), 19–24.
3. Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W. & Taylor, J. G. Emotion recognition in human-computer interaction, *IEEE Signal Processing Magazine*, 18, 1(2001), 32-80.
4. Cowie, R., Gunes, H., McKeown, G., Vaclau-Schneider, L. , Armstrong, J. & Douglas-Cowie, E. The emotional and communicative significance of head nods and shakes in a naturalistic database, *Proc. of LREC Int. Workshop on Emotion* (2010), 42-46.
5. Ekman, P. & Friesen, W.V. Unmasking the face: a guide to recognizing emotions from facial clues (2003) Cambridge, MA.
6. Ekman, P. & Friesen, W. V., Head and body cues in the judgment of emotion: A reformulation, *Perceptual and Motor Skills*, 24 (1967), 711-724
7. Fontaine, J. R., Scherer, K. R., Roesch, E. B. & Ellsworth, P. The world of emotion is not two-dimensional, *Psychological Science*, 18 (2007), 1050-1057.
8. Frijda, N. H. (1986) *The emotions*, Cambridge University Press.
9. Gokcay, D. & Yildirim, G. (Eds.) (2010) Affective Computing and Interaction: Psychological, Cognitive and Neuroscientific Perspectives, IGI Global.
10. Grandjean, D., Sander, D. & Scherer, K. R. Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization, *Consciousness and Cognition*, 17, 2 (2008), 484-495.
11. Gunes, H. & Pantic, M. Automatic, dimensional and continuous emotion recognition, *Int. Journal of Synthetic Emotions*, 1, (1) 2010, 68-99.
12. Gunes, H. Piccardi, M. & Pantic, M. From the lab to the real world: Affect recognition using multiple cues and modalities, *In Affective Computing, Focus on Emotion Expression, Synthesis and Recognition* (2008), 185-218.
13. Haag, A., Goronzy, S., Schaich, P. & Williams, J. Emotion recognition using bio-sensors: First steps towards an automatic system, *LNCS 3068* (2004), 36-48.
14. Kanluan, I., Grimm, M. & Kroschel, K. Audio-visual emotion recognition using an emotion recognition space concept, *Proc. of the 16th European Signal Processing Conference* (2008).
15. Kim, J., Bimodal emotion recognition using speech and physiological changes, *In Robust Speech Recognition and Understanding*, (2007) 265-280.
16. Kleinsmith, A. & Bianchi-Berthouze, N., Recognizing affective dimensions from body posture, *LNCS 4738* (2007), 48-58.
17. Kulic, D. & Croft, E. A. Affective state estimation for human–robot Interaction, *IEEE Tran. on Robotics*, 23, 5(2007), 991- 1000.
18. Nicolaou, M. A., Gunes H. & Pantic, M. Automatic segmentation of spontaneous data using dimensional labels from multiple coders, *Proc. of LREC Int. Workshop on Multimodal Corpora* (2010), 43-48.
19. Ortony, A., Clore, G. L. & Collins, A. (1988) *The cognitive structure of emotions*, Cambridge Univ. Press.
20. Picard, R.W., Emotion research by the people, for the people, *Emotion Review* (2010), in press.
21. Picard, R.W., Future affective technology for autism and emotion communication, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 1535 (2009), 3575-3584
22. Pun, T., Alecu, T.I., Chanel, G., Kronegg, J. & Voloshynovskiy, S. Brain-computer interaction research at the Computer Vision and Multimedia Laboratory, University of Geneva, *IEEE Tran. on Neural Systems and Rehabilitation Engineering*,. 14, 2 (2006), 210-213.
23. Russell, J. A. A circumplex model of affect, *J. of Personality and Social Psychology*, 39 (1980)., 1161-1178.
24. Scherer, K.R., Schorr, A., Johnstone, T. (Eds.). (2001) *Appraisal Processes in Emotion: Theory, Methods, Research,* Oxford Univ. Press.
25. Schröder, M., Bevacqua, E., Eyben, F., Gunes, H., Heylen, D., Maat, M., Pammi, S., Pantic, M., Pelachaud, C., Schuller, B., Sevin, E., Valstar, M. & Wöllmer, M. A demonstration of audiovisual sensitive artificial listeners, *Proc. of Int. Conf. on Affective Computing and Intelligent Interaction* (2009), 1, 263-264.
26. Tsiamyrtzis, P., Dowdall, J., Shastri, D., Pavlidis, I. & Frank, M.G. Imaging facial physiology for the detection of deceit, *Int. J. of Computer Vision*, 71, 2 (2007), 197-214.
27. Wöllmer, M., Eyben, F., Reiter, S., Schuller, B., Cox, C., Douglas-Cowie, E. & Cowie, R. Abandoning emotion classes - Towards continuous emotion recognition with modelling of long-range dependencies, *Proc. of Interspeech* (2008), 597-600.

Proceedings of

Measuring Behavior 2010

7th International Conference on Methods and Techniques in Behavioral Research

Eindhoven, The Netherlands
August 24-27, 2010

Editors
Andrew Spink
Fabrizio Grieco
Olga Krips
Leanne Loijens
Lucas Noldus
Patrick Zimmerman

Noldus
Information Technology

2010

Bibliographical reference

Proceedings of Measuring Behavior 2010, 7th International Conference on Methods and Techniques in Behavioral Research (Eindhoven, The Netherlands, August 24-27, 2010)

Edited by A.J. Spink, F. Grieco, O.E. Krips, L.W.S. Loijens, L.P.J.J. Noldus, and P.H. Zimmerman

ISBN 978-90-74821-86-5

Copyright

© 2010 Noldus Information Technology bv

P.O. Box 268

6700 AG Wageningen

The Netherlands

www.noldus.com

Printed in the Netherlands

No part of this publication may be reproduced, transcribed, stored in a retrieval system or translated into any other language, in whole or in part, in any form or by any means, without the prior written permission of the publisher.

Some of the product names referred to in this book are registered trademarks or proprietary names, even though specific reference to this fact is not always made in the text. Therefore, the appearance of a name without designation as proprietary is not to be considered as a representation by the publisher that it is in the public domain.

Table of Contents

Keynote Lectures

Symposia

Symposium 1: Improving Sport Performance

Symposium 2: Large and Small Scale Physiological Recordings in Behavioural Context

Symposium 6: Unveiling Affective Signals

Full papers

Full paper session 1: Measuring the Behavior of Laboratory Animals

Welcome to Measuring Behavior 2010

It is my great pleasure to welcome you to Measuring Behavior 2010, the 7th International Conference on Methods and Techniques in Behavioral Research. This conference edition is hosted at the High Tech Campus in Eindhoven. With over 90 companies and institutes, the HTC brings together a dynamic mix of multinational companies, small and medium-sized businesses and technology start-up companies. Campus residents share knowledge, experience, open laboratories and technical infrastructure, enabling better, faster and more cost efficient innovation. An open environment that fuels opportunities for valuable R&D, for successful business partnerships.

In this year's conference we bring closer together a diversity of communities ranging from neuroscience and zoology to psychology and consumer behavior. The conference location could not have been more symbolic for meeting with a multidisciplinary community interested in sharing methods and techniques for conducting behavioral research.

This year's Measuring Behavior conference features a very strong technical program, assembled under the expert leadership of Program Co-Chairs Emilia Barakova and Andrew Spink. Together with the Scientific Program Committee and expert reviewers from the community, they undertook the difficult job of carefully evaluating the large number of submitted papers, considering the merits of each through detailed reviews and selecting a technical program of the highest caliber. With a technical program bringing symposia, paper sessions, demonstrations, tutorials, user meetings, workshops, scientific tours and exhibitions there offer the best setting for productive cross-fertilization between research fields in the area of measuring behavior.

Again, welcome to Measuring Behavior 2010 at the High Tech Campus. I wish you a very productive and informative conference and hope that you will take the opportunity to strengthen your network with the Measuring Behavior community.

Boris de Ruyter
Measuring Behavior 2010 Conference Chair

The *Measuring Behavior* Conferences

Measuring Behavior is a unique conference about methods and techniques in behavioral research. While most conferences focus on a specific domain, *Measuring Behavior* creates bridges between disciplines by bringing together people who may otherwise be unlikely to meet each other. At a *Measuring Behavior* meeting, you find yourself among ethologists, behavioral ecologists, neuroscientists, experimental psychologists, human factors researchers, movement scientists, robotics engineers, software designers, human-computer interaction specialists… to mention just a few. While the research questions and applications may be highly diverse, all delegates share an interest in methods, techniques and tools for studying behavior. Experience tells us that the focus on methodological and technical themes can lead to a very productive cross-fertilization between research fields. Crossing the boundaries between disciplines and species (from insects to astronauts) can be extremely inspiring. For many delegates, attending a *Measuring Behavior* meeting is an eye-opening experience, to find out which interesting (and often highly relevant) developments are taking place in domains they usually don't venture into.

Measuring Behavior started in 1996 as a workshop in the framework of a European research project "Automatic Recording and Analysis of Behavior", aimed at sharing the results of our project with colleagues from abroad. Organized by Noldus Information Technology and hosted by Utrecht University, *Measuring Behavior '96* attracted over 150 participants from 25 countries. Encouraged by the international interest, it was decided to make *Measuring Behavior* a recurring conference. In the years that followed, the conference travelled to six other Dutch university towns: Groningen (1998), Nijmegen (2000), Amsterdam (2002), Wageningen (2005), Maastricht (2008) and now Eindhoven (2010).

Over the years, *Measuring Behavior* has developed a formula with a mix of ingredients that has proven quite successful. The meeting is always held in a university town where research on human or animal behavior is prominent, with local scientists playing a prominent role in the conference organization (see table below). Noldus Information Technology serves as conference organizer and main sponsor. For a small company like ours, the conference is a major investment. The registration fees just cover the direct expenses associated with the meeting; the hours spent on the organization (several person-years) are on our account. We gladly do this, because we believe that the focused attention on behavior research methods and techniques will eventually lead to a higher demand for our tools. To prevent commercial bias, however, the scientific program is put together under auspices of an independent Scientific Program Committee, consisting of international experts from a broad variety of disciplines (see the Scientific Program Committee on page 501). We are very grateful for their effort to review papers and the helpful input during email exchanges as well as all the other reviewers of the papers in the scientific program.

Year	**City**	**Conference chair**
1996	Utrecht	Berry Spruijt
1998	Groningen	Jaap Koolhaas
2000	Nijmegen	Alexander Cools
2002	Amsterdam	Gerrit van der Veer
2005	Wageningen	Louise Vet
2008	Maastricht	Harry Steinbusch
2010	Eindhoven	Boris de Ruyter

Over the years, the conference has grown significantly in size, from 153 delegates in 1996 to more than 400 in 2008. At this size, the event is large enough to cover a wide range of topics, yet still small enough for a social program with all delegates. *Measuring Behavior* has also become a truly global meeting: delegates come from dozens of countries on all continents. 2010 is the most international version so far with participants from more than 35 countries and for the papers submitted individually the host country is not longer the one with the most papers (that honor goes to the USA).

In the scientific program, well balanced between human and animal research, one finds a variety of formats for presentation, interaction and exchange of information. The traditional oral papers (full papers) and poster presentations have always been central to the conference. Increasingly, special symposia– focusing on a current methodological or technical theme – are proposed by experts from various disciplines. These symposia illustrate the widening scope of the conference as well as trends in science. From the start, there have been symposia on topics in behavioral neuroscience (such as animal models for human disease or automatic behavior recognition in rats and mice) and data analysis and statistics (such as sequential analysis and pattern detection). Subsequently, the scope extended towards psychology, human factors, ergonomics and movement science. And at this year's conference we see the arrival of novel topics such as human-robot interaction, measuring behavior in the operating theatre and behavior of forensic scientists.

Besides oral presentations in symposia or free paper sessions, the conference program always includes ample time for posters and demonstrations of software or equipment by participants. The latter are actively encouraged, because it is a format not supported by most other conferences. Full demonstrations in a seminar room were new at the previous Measuring Behavior and we are pleased to see that has really taken off, with 12 separate demonstrations on the Thursday afternoon. For several academic inventors, the presentation of their prototype software or hardware tool at *Measuring Behavior* paved the way towards commercialization. This is how CatWalk and FaceReader found their way to the Noldus product portfolio. We hope that scientists will continue to present inventions at *Measuring Behavior* and discuss commercialization opportunities with the vendors present at the meeting.

Another attractive element of the conference is the scientific tours, guided visits to behavioral research facilities and laboratories in and around the hosting university. Tutorials, short courses – mostly about software tools and instruments – taught by expert instructors, have also become a popular program element. Other program elements are user meetings (organized by manufacturers of research tools), and workshops. At this year's conference, the latter have become more prominent: seven workshops are being held, about topics ranging from autism research to GPS tracking and behavior recognition in wildlife. Finally, there is the commercial exhibition of scientific instruments and software related to behavioral research.

Measuring Behavior is a scientific conference, so special attention is paid to publication of the work presented at the meeting. We started off with a program book and an abstracts book. In 2005 we added printed conference proceedings with short papers and a conference CD. Because of the overlap between the abstracts and proceedings books, we have gone back to two books: a program book and printed proceedings of extended abstracts (short papers). Then there is the conference website (www.measuringbehavior.org). After each conference, the Measuring Behavior website is converted into an archival site, with abstracts of all presentations, which remain accessible. The websites of past conferences form a valuable resource on methods and techniques for behavioral research. Presenting authors will have noticed that they were asked to submit using a rather complex template. This was to enable us to archive all the reviewed papers to the Digital Library of the ACM (the Association for Computing Machinery), which is an important publication channel for those studying human-computer interaction. Finally, selected presentations will be published as full papers in the Journal of Integrative Neuroscience.

Now you find yourself at the 7th *Measuring Behavior* conference. The organizers have done their best to prepare an optimal mix of scientific, technical, social and culinary ingredients. We hope that you will find *Measuring Behavior 2010* a rewarding experience and wish you a pleasant stay in Eindhoven.

Lucas P.J.J. Noldus

Managing Director, Noldus Information Technology BV

l.noldus@noldus.nl