# Towards Emotion Recognition in Human Computer Interaction

Günther Palm[1,2,⋆] and Michael Glodek[2]

[1] Domenico Labate, MECMAT - Mediterranea University of Reggio Calabria, via Graziella Feo di Vito, 89122 Reggio Calabria, Italy
[2] Institute of Neural Information Processing, University of Ulm, Germany
firstname.lastname@uni-ulm.de

**Abstract.** The recognition of human emotions by technical systems is regarded as a problem of pattern recognition. Here methods of machine learning are employed which require substantial amounts of 'emotionally labeled' data, because model based approaches are not available. Problems of emotion recognition are discussed from this point of view, focusing on problems of data gathering and also touching upon modeling of emotions and machine learning aspects.

**Keywords:** Affective computing, emotion recognition, human computer interaction, companion systems.

## 1 Introduction

In the field of human-computer interaction (HCI) there is a growing interest in using not only explicit function-oriented signals for the cooperation between a human user and a technical system, but also implicit "affective" or "emotional" signals.

In spite of a long tradition in psychology (e.g. [1,2,3,4,5,6,7,8]) there still is no undisputed definition of emotions, neither of particular emotions nor of the general concept of emotion. From a biological perspective emotions can be described as modulations of animal behavior in response to certain extreme circumstances. For example, in extreme danger, an animal may focus its attention to the relevant sensory modalities and to particular spatio-temporal constellations in these modalities; in addition it may mobilize all its physical energy and strength in particular groups of muscles to enable quick reactions like fleeing or fighting, and it may shift its evaluative weights (which will be explained in a bit more detail below) towards taking more risks of injuries and pain in the course of actions. This scenario describes a modulation of the three main ingredients of goal-directed behavior: sensation, action and evaluation. It is plausible that in the course of evolution and also of individual development animals have found

⋆ Corresponding author.

B. Apolloni et al. (Eds.): *Neural Nets and Surroundings*, SIST 19, pp. 323–336.
DOI: 10.1007/978-3-642-35467-0_32 

a small number of such combined modulatory settings that are useful for the modulation of behavior in certain situations. These settings may be what we call emotions.

In social animals it may also be useful to signal these settings to other animals in order to prevent, for example, that animals of the same species hurt each other unnecessarily. So emotional states may also be conveyed by social signals, which in humans often are unconsciously produced in addition to or as modification of our explicit communication.

The biological approach sketched here may sound provocative to some psychologists, but it boils down to defining emotions as particular modulators of behavior which may also have some physiological consequences, for example concerning heart-rate, blood-flow, body-heat and skin resistance.

Some attempts have been made to define "basic" emotion categories (typically around six of them, see for example [9], but the varying degree of similarity between these categories suggests an embedding of them into a 2- or 3-dimensional space; most commonly in a circular fashion ([3] or the "emotion wheel" of [8]). The dimensions or axes of this space have also been identified ([10,6]) as "value" (from negative or "bad" to positive or "good"), "activation" (from low or "calm" to high or "agitated") and "dominance" or perhaps also "competence" or "influence" (from low or "passive", "enduring" to high or "active", "in control"). Another aspect of the classification of emotions is their variable strength or expressiveness.

When we try to incorporate emotionality in HCI, we also have to consider emotions from the computer science point-of-view. Here there are three tasks that can be distinguished although they are often used in combination:

1. detection and classification of human emotions,
2. modeling of emotions in an artificial agent,
3. displaying emotional signals towards the human user.

1. *Emotion detection* can be viewed as a particular branch of pattern recognition and will be in the focus of this article.

2. *Emotion modeling* is an interesting enterprise which may follow the biological approach described above and thereby relate to neurobiological and physiological research on this topic (e.g. [11,12,13,14,15,16]) The neuroscientific accounts of emotions are concentrating on fMRI experiments showing emotional reactions in particular brain areas and also on the idea of explaining the cognitive faculty of empathy as a use of the own "emotion generating system" in those areas to simulate or "mirror" the emotions of others ([17,18,19,20,21,22,23,24]), similar to the mirror neuron system representing intentional actions. Theoretical approaches often rely on variations of reinforcement learning ([25,26,27,12,28,29,30,31,32,33,15,34]). The basic idea of reinforcement learning (RL) is that an agent learns to predict an evaluative signal (called reward) based on its sensory input and the course of action it is about to take. It then takes the course of action that maximizes the predicted reward. Biologically realistic versions of RL assume a small number

of different motives or objectives like eating, drinking, having sex, exploring territory, avoiding pain, which the agent may want to optimize with different weights or different priorities depending on its current situation (these are the evaluative weights mentioned above).
There is also an interesting modeling approach by D. Dörner and his scholars ([35,36,37,38]) that is closely related to these ideas. Another approach to the modeling of emotions is motivated from cognitive psychology ([39,40,41] ) treating the relation between emotions and motivations or appraisals on a higher cognitive level with less emphasis on the underlying physiological processes and on the dynamical temporal aspects of emotions.

3. *Emotion display* has become quite popular in HCI in applications concerned with entertainment and computer games. It makes use of the fact that humans readily project emotions into artificial agents they interact with. In a feedback situation of HCI it can be useful to combine emotion recognition with emotion display, for example in order to improve and stabilize the emotion recognition performance on both sides in a dyadic HCI.

In fact, one can perhaps divide the whole area of affective computing into those domains where the expression of emotions by the computer is more important, which will most likely occur in recreational activities like entertainment, edutainment, or game playing, and those domains where the recognition of human emotions by the computer is more important, which will most likely occur in the context of work, when the computer helps to achieve a given task. In our collaborative research center (see `http://www.sfb-trr-62.de/`) we focus on the second type of interaction and we define *Companion technology* as the technology that helps to improve this kind of HCI.

## 2 Recognition of Emotions or User Dispositions

In HCI the recognition of emotions has turned out to be a very hard problem of pattern recognition and machine learning due to the following reasons:

- Emotions are rare. They do not occur often in human interaction and perhaps even less in HCI.
- Most emotions can occur in different degrees and they occur only in weak degrees most of the time.
- This often leads to a high degree of uncertainty about the recognized emotion and, in addition, it may be necessary to distinguish the degree of recognition uncertainty from the degree of expressiveness of the present emotion.
- Emotions are expressed multi-modally. So it is often useful and even necessary to combine several sensory modalities (audition, vision, and biophysical measurements, if available) in their detection and classification.
- Emotion recognition is context dependent. Humans often cannot correctly identify emotions unless they are provided within broader context of the given situation. For artificial systems this implies that one should make use

**Table 1.** Multimodal datasets of HCI with annotated emotions

| *Name* | *Modalities*[1] | *Size*[2] | *Annotation*[3] | *HCI type*[4] | *Remarks*[5] | *Reference* |
|---|---|---|---|---|---|---|
| AVEC | A,V | 95/13/7.5h | 4 D<br>2-8 raters | D | T,E | [78] |
| EmoRec | A,V,P | 110/110/73h | 3 D<br>not req. | S,T | E | [75] |
| Humaine | A,V | 50/?/3.3h | 4 D<br>n/a | D | E | [79] |
| Last Minute | A,V,P | 126/126/63h | not yet<br>annotated | S,C | T,C | [80] |
| Nimitek | A,V | 10/10/15h | 6 C<br>n/a | T | T,E | [81] |
| PIT | A,V | 37/74/9h | > 9 C<br>2-3 raters | C,O | C | [82] |
| SAL | A,V | 20/4/10h | 6 C, 4 D<br>n/a | D | E | [83] |
| Smartkom | A,V | 224/224/16.8h | 9 C<br>n/a | C,O | T | [84] |

[1] (A) audio, (P) physiological, (V) video
[2] #/#/# number of recordings/number of subjects/hours of recording
[3] # number of labels, (C) categories (D) dimensions;
number of raters, (n/a) number of raters not available,
(not req.) rating not required
[4] (T) trainer/teacher, (M) monitor, (O) organizer, (C) consultant, (S) servant,
(D) discourse
[5] (T) transcript available, (E) emotions are evoked, (C) multiple cameras

of high-level symbolic information, if available, and at least use temporal integration of the given multimodal signals over a longer time (e.g. covering a whole interaction sequence, not just one utterance or turn in a dialogue).

These observations have been made in our own attempts to create a "*Companion technology*" ([42] ) for HCI that allows the computer to react to human emotional signals. There is not much scientific literature yet on these topics to substantiate these preliminary observations (e.g. [43,44,45,46,47,48,49,50]), since the field of "affective computing" is still in its infancy. Early experiments on unimodal emotion recognition ([51,7,52,53,54,55]) had to rely on artificial acted emotions. And even on these data (e.g. [56,57,58,59]) the typical performance for 4 or 5 emotion categories was around 70 to 80 percent, usually about as bad or just a little worse than human performance on the same data.

There are some projects providing large multimodal databases of almost natural human conversations (see [60,61,62,63,64,65,66,67,68,69,70]) which can be used to substantiate the 5 points above, but not much work on natural emotion recognition has appeared so far (e.g. [48,71,49,72,73] [74,75,76,77]).

A serious problem of emotion recognition from the engineering point of view is that, strictly speaking, there is no ground truth that can be used for labeling. Even if one tries to work with induced emotions the exact emotional category that is induced may strongly depend on the character of the subject, in addition to ethical problems involved in the induction of strong negative emotions. The only available 'gold-standard' is human rating of the observed emotional behavior and this is far from being unequivocal.

Concerning emotions in HCI the available data are still very sparse. Currently we are aware just of a few reasonably extensive multimodal HCI databases (see Table 1). In fact, any project of data-collection for affective HCI is faced with a number of very particular challenges:

– Technical problems with gathering and displaying large amounts of synchronized multi-modal data in a reasonable HCI scenario (often requiring WOZ methodology).
– Determining the relevant label categories for the data, which may even be application- or scenario-dependent (see below).
– Developing tools that make it possible for annotators to add (un)certainty values to their labels (like [85,86,87]).

### 2.1 Identifying User Dispositions in Companion Technology

In spite of these problems we expect much further progress in affective HCI for the development of *Companion* systems, based on a few conceptual and technical developments in the next years.

First, we said that emotions are rare in HCI, and indeed some of the genuine human emotions may not even be relevant for typical *Companion* systems. Hence one can try to restrict the emotional categories that have to be recognized and distinguished to the practically relevant ones. To this end it would of course be useful not to consider every application separately, but to identify a small set of typical applications or *Companion tasks*. In our understanding of *Companion* systems we are not focusing on entertainment systems, but rather on applications where the *Companion* system helps the user to achieve some goal. A typical problem of this kind occurs if the user has to rely on a conventional technical system which is complicated or partially unknown to the user. In such a case the *Companion* system can mediate between the user and the conventional system. Another type of application occurs in training or teaching where the system guides the user in extending his knowledge or his practical or physical capabilities.

To arrive at a short list of generic *Companion* tasks we can consider the typical examples of assistance that are proposed or already offered by modern technical systems in the household (cleaning, cooking, preparing or organizing meals, invitations), in the car (navigation, driving assistance), at work (keeping track of dates, meetings and duties, finding information, e.g. on products, transportation, addresses, making connections to other people), simplifying or personalizing use of machinery (cameras, coffee-machines, cell-phones, audio-video-displays, diagnosis and repair), in education (learning, edutainment), in buying, ordering things, in the hospital and rehabilitation (monitoring, control of medication, motivation of physical exercises).

We have attempted to condense this down to the following five types of *Companions*: **Trainer/teacher**, **monitor**, **organizer**, **consultant**, **servant**.

When we now consider the emotional dispositions of the user that will be relevant for the reactions of the *Companion* in these cases, it turns out that in practically all cases it is the disposition or attitude of the user towards the task at hand and towards the *Companion* system that matters. The central affective goal of the system is to maintain a positive attitude of the user, for example by adjusting the amount, redundancy and complexity of the information it is providing, or by giving ensuring or just evaluating feedback to the user (and, of course, by functioning properly). Thus the main practical objective in emotion recognition for such systems will be to distinguish different kinds of **negative** emotions or attitudes from each other, and also from neutral or “no problem”. Here we propose to distinguish the following seven: **Bored**, **disengaged**, **frustrated**, **helpless**, **over-strained**, **angry**, **impatient**.

Again this short list of specific application oriented user attitudes or dispositions has to be regarded as preliminary and it certainly needs further explanation, for example in terms of emotion dimensions like activation and dominance, in particular when it is used for labeling. Such a short list, however, would help a lot to simplify the task of emotion recognition for *Companion* systems and to unify the labeling of corresponding affective HCI databases for *Companion* systems.

Another reason for anticipating a fast progress in emotion recognition lies in the development of new ideas and techniques in pattern recognition and machine learning. There are a number of new ideas concerning information and classifier fusion ([88,89,90,91,92,93,94,95,96,97]) which have lead to better theoretical understanding and practical results in the fusion of information from different sources and have already been applied to the recognition of user dispositions or emotions (e.g. [96,98,99,100,101,102,103]) There are also new results on semi-supervised learning ([104,105,106,107,108,109,110,111,112,113,114,115,116] [117,118]) and on learning from uncertain teacher signals ([119,120,121]) that can be used to simplify the labeling of large datasets with uncertain or partially missing labels.

We believe that, based on these ideas and developments, the field of affective computing for *Companion* systems and also for HCI in general is going to make considerable progress in the next years.

**Acknowledgment.** This work was stimulated by extensive discussions and financially supported within the Transregional Collaborative Research Centre SFB/TRR 62 ”Companion-Technology for Cognitive Technical Systems” funded by the German Research Foundation (DFG).

## References

1. Cannon, W.: The James-Lange theory of emotions: A critical examination and an alternative theory. The American Journal of Psychology 39(1/4), 106–124 (1927)
2. Ekman, P., Friesen, W.: The repertoire of nonverbal behavior: Categories, origins, usage, and coding. Semiotica 1(1), 49–98 (1969)

3. Russell, J.: A circumplex model of affect. Journal of Personality and Social Psychology 39(6), 1161 (1980)
4. Ekman, P.: Facial expression and emotion. American Psychologist 48, 384–392 (1993)
5. Picard, R.W.: Affective Computing. The MIT Press (1997)
6. Russell, J., Barrett, L.: Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. Journal of Personality and Social Psychology 76(5), 805 (1999)
7. Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., Taylor, J.: Emotion recognition in human-computer interaction. IEE Signal Processing Magazine 18(1), 32–80 (2001)
8. Scherer, K.: What are emotions? And how can they be measured? Social Science Information 44(4), 695–729 (2005)
9. Ekman, P.: An argument for basic emotions. Cognition & Emotion 6(3-4), 169–200 (1992)
10. Mehrabian, A.: Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. Current Psychology 14(4), 261–292 (1996)
11. Damasio, A.: Descartes' error and the future of human life. Scientific American 271(4), 144–145 (1994)
12. Dayan, P., Balleine, B.: Reward, motivation, and reinforcement learning. Neuron 36(2), 285–298 (2002)
13. Taylor, J., Fragopanagos, N.: The interaction of attention and emotion. Journal of Neural Networks 18(4), 353–369 (2005)
14. Fogassi, L., Ferrari, P.: Mirror neurons and the evolution of embodied language. Current Directions in Psychological Science 16(3), 136–141 (2007)
15. Dayan, P., Huys, Q.: Serotonin in affective control. Annual Review of Neuroscience 32, 95–126 (2009)
16. Niv, Y.: Reinforcement learning in the brain. Journal of Mathematical Psychology 53(3), 139–154 (2009)
17. Rizzolatti, G., Arbib, M.: Language within our grasp. Journal of Trends in Neurosciences 21(5), 188–194 (1998)
18. Rizzolatti, G., Fogassi, L., Gallese, V.: Mirrors in the mind. Scientific American 295(5), 54–61 (2006)
19. Rizzolatti, G., Sinigaglia, C.: Mirrors in the brain: How our minds share actions, emotions. Oxford University Press (2008)
20. Arbib, M.: Beyond the Mirror: Biology and Culture in the Evolution of Brain and Language. Oxford University Press (2005)
21. Arbib, M.: Action to language via the mirror neuron system. Cambridge University Press (2006)
22. Bonaiuto, J., Rosta, E., Arbib, M., et al.: Extending the mirror neuron system model, I. Audible actions and invisible grasps. Journal of Biological Cybernetics 96(1), 9 (2007)
23. Bonaiuto, J., Arbib, M.: Extending the mirror neuron system model, II: What did i just do? A new role for mirror neurons. Journal of Biological Cybernetics 102(4), 341–359 (2010)
24. Gallese, V., Goldman, A.: Mirror neurons and the simulation theory of mind-reading. Journal of Trends in Cognitive Sciences 2(12), 493–501 (1998)

25. Albus, J.: Outline for a theory of intelligence. IEEE Transactions on Systems, Man and Cybernetics 21(3), 473–509 (1991)
26. Bertsekas, D., Tsitsiklis, J.: Neuro-dynamic programming. Journal of Optimization and Neural Computation 3 (1996)
27. Sutton, R., Barto, A.: Reinforcement learning: An introduction, vol. 1. Cambridge University Press (1998)
28. Wörgötter, F.: Actor-critic models of animal control — A critique of reinforcement learning. In: Proceedings of International Symposium on Engineering of Intelligent Systems, EIS (2004)
29. Wörgötter, F., Porr, B.: Temporal sequence learning, prediction, and control: A review of different models and their relation to biological mechanisms. Journal of Neural Computation 17(2), 245–319 (2005)
30. Oudeyer, P., Kaplan, F., Hafner, V.: Intrinsic motivation systems for autonomous mental development. IEEE Transactions on Evolutionary Computation 11(2), 265–286 (2007)
31. Izhikevich, E.: Solving the distal reward problem through linkage of STDP and dopamine signaling. Journal of Cerebral Cortex 17(10), 2443–2452 (2007)
32. Botvinick, M., Niv, Y., Barto, A.: Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. Journal of Cognition 113(3), 262–280 (2009)
33. Bhatnagar, S., Sutton, R., Ghavamzadeh, M., Lee, M.: Natural actor–critic algorithms. Automatica 45(11), 2471–2482 (2009)
34. Huys, Q., Dayan, P.: A Bayesian formulation of behavioral control. Journal of Cognition 113(3), 314–328 (2009)
35. Bartl, C., Dörner, D.: PSI: A theory of the integration of cognition, emotion and motivation. In: Proceedings of the European Conference on Cognitive Modelling (ECCM), DTIC Document, pp. 66–73 (1998)
36. Dörner, D.: Bauplan fur eine Seele (Blueprint of a Soul). RoRoRo (1999)
37. Hille, K.: Synthesizing emotional behavior in a simple animated character. Journal of Artificial Life 7(3), 303–313 (2001)
38. Bach, J.: Principles of synthetic intelligence — PSI: an architecture of motivated cognition. Oxford University Press (2008)
39. Ortony, A., Clore, G., Collins, A.: The cognitive structure of emotions. Cambridge University Press (1988)
40. Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H., et al.: Understanding and sharing intentions: The origins of cultural cognition. Behavioral and Brain Sciences 28(5), 675–690 (2005)
41. Marsella, S., Gratch, J.: EMA: A process model of appraisal dynamics. Journal of Cognitive Systems Research 10(1), 70–90 (2009)
42. Wendemuth, A., Biundo, S.: A Companion Technology for Cognitive Technical Systems. In: Esposito, A., Esposito, A.M., Vinciarelli, A., Hoffmann, R., Müller, V. (eds.) Cognitive Behavioural Systems 2011. LNCS, vol. 7403, pp. 89–103. Springer, Heidelberg (2012)
43. Scherer, K., Ceschi, G.: Criteria for emotion recognition from verbal and nonverbal expression: studying baggage loss in the airport. Personality and Social Psychology Bulletin 26(3), 327–339 (2000)

44. Batliner, A., Fischer, K., Huber, R., Spilker, J., Nöth, E.: How to find trouble in communication. Journal of Speech Communication 40(1), 117–143 (2003)
45. Batliner, A., Zeißler, V., Frank, C., Adelhardt, J., Shi, R., Nöth, E.: We are not amused-but how do you know? User states in a multi-modal dialogue system. In: Proceedings of the European Conference on Speech Communication and Technology, Eurospeech (2003)
46. Batliner, A., Hacker, C., Steidl, S., Nöth, E., Haas, J.: From Emotion to Interaction: Lessons from Real Human-Machine-Dialogues. In: André, E., Dybkjær, L., Minker, W., Heisterkamp, P. (eds.) ADS 2004. LNCS (LNAI), vol. 3068, pp. 1–12. Springer, Heidelberg (2004)
47. Esposito, A.: The amount of information on emotional states conveyed by the verbal and nonverbal channels: Some perceptual data. Journal of Progress in Nonlinear Speech Processing, 249–268 (2007)
48. Schuller, B., Seppi, D., Batliner, A., Maier, A., Steidl, S.: Towards more reality in the recognition of emotional speech. In: Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 4, pp. 941–944. IEEE (2007)
49. Wendt, C., Popp, M., Karg, M., Kuhnlenz, K.: Physiology and HRI: Recognition of over-and underchallenge. In: Proceedings of the Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 448–452. IEEE (2008)
50. Batliner, A., Steidl, S., Schuller, B., Seppi, D., Vogt, T., Wagner, J., Devillers, L., Vidrascu, L., Aharonson, V., Kessous, L., et al.: Whodunnit-searching for the most important feature types signalling emotion-related user states in speech. Journal of Computer Speech & Language 25(1), 4–28 (2011)
51. Tian, Y., Kanade, T., Cohn, J.: Recognizing action units for facial expression analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(2), 97–115 (2001)
52. Lee, C., Yildirim, S., Bulut, M., Kazemzadeh, A., Busso, C., Deng, Z., Lee, S., Narayanan, S.: Emotion recognition based on phoneme classes. In: Proceedings of the Annual Conference of the International Speech Communication Association (ISCA), Interspeech (2004)
53. Wagner, J., Vogt, T., André, E.: A Systematic Comparison of Different HMM Designs for Emotion Recognition from Acted and Spontaneous Speech. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 114–125. Springer, Heidelberg (2007)
54. Vlasenko, B., Schuller, B., Wendemuth, A., Rigoll, G.: Frame vs. Turn-Level: Emotion Recognition from Speech Considering Static and Dynamic Processing. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 139–147. Springer, Heidelberg (2007)
55. Maganti, H.K., Scherer, S., Palm, G.: A Novel Feature for Emotion Recognition in Voice Based Applications. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 710–711. Springer, Heidelberg (2007)
56. Kanade, T., Cohn, J., Tian, Y.: Comprehensive database for facial expression analysis. In: Proceedings of the International Conference on Automatic Face and Gesture Recognition (FG), pp. 46–53. IEEE (2000)
57. Wendt, B., Scheich, H.: The "Magdeburger Prosodie-Korpus". In: Proceedings of the International Conference on Speech Prosody, pp. 699–701 (2002)

58. Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., Weiss, B.: A database of German emotional speech. In: Proceedings of the European Conference on Speech Communication and Technology, Eurospeech (2005)
59. Bänziger, T., Scherer, K.R.: Using Actor Portrayals to Systematically Study Multimodal Emotion Expression: The GEMEP Corpus. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 476–487. Springer, Heidelberg (2007)
60. Douglas-Cowie, E., Campbell, N., Cowie, R., Roach, P.: Emotional speech: Towards a new generation of databases. Journal of Speech Communication 40(1), 33–60 (2003)
61. Douglas-Cowie, E., Cowie, R., Sneddon, I., Cox, C., Lowry, O., McRorie, M., Martin, J.-C., Devillers, L., Abrilian, S., Batliner, A., Amir, N., Karpouzis, K.: The HUMAINE Database: Addressing the Collection and Annotation of Naturalistic and Induced Emotional Data. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 488–500. Springer, Heidelberg (2007)
62. Campbell, N.: Databases of expressive speech. Journal of Chinese Language and Computing 14(4), 295–304 (2004)
63. Campbell, N.: Listening between the lines: A study of paralinguistic information carried by tone-of-voice. In: Proceedings of the International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages (2004)
64. Campbell, N., Sadanobu, T., Imura, M., Iwahashi, N., Noriko, S., Douxchamps, D.: A multimedia database of meetings and informal interactions for tracking participant involvement and discourse flow. In: Proceedings of the International Conference on Language Resources and Evaluation, LREC (2006)
65. Cowie, R., Douglas-Cowie, E., Cox, C.: Beyond emotion archetypes: Databases for emotion modelling using neural networks. Journal of Neural Networks 18(4), 371–388 (2005)
66. Martin, O., Kotsia, I., Macq, B., Pitas, I.: The eNTERFACE'05 audio-visual emotion database. In: Proceedings of the International Conference on Data Engineering Workshops (ICDE), p. 8. IEEE (2006)
67. Grimm, M., Kroschel, K., Narayanan, S.: The Vera am Mittag German audio-visual emotional speech database. In: Proceedings of the International Conference on Multimedia and Expo (ICME), pp. 865–868. IEEE (2008)
68. Strauß, P.M., Hoffmann, H., Minker, W., Neumann, H., Palm, G., Scherer, S., Traue, H., Weidenbacher, U.: The PIT corpus of german multi-party dialogues. In: Proceedings of the International Conference on Language Resources and Evaluation, LREC (2008)
69. McKeown, G., Valstar, M., Cowie, R., Pantic, M.: The SEMAINE corpus of emotionally coloured character interactions. In: Proceedings of the International Conference on Multimedia and Expo (ICME), pp. 1079–1084. IEEE (2010)
70. Oertel, C., Scherer, S., Wagner, P., Campbell, N.: On the use of multimodal cues for the prediction of involvement in spontaneous conversation. In: Proceedings of the Annual Conference of the International Speech Communication Association (ISCA), Interspeech, pp. 1541–1544. ISCA (2011)
71. Caridakis, G., Castellano, G., Kessous, L., Raouzaiou, A., Malatesta, L., Asteriadis, S., Karpouzis, K.: Multimodal Emotion Recognition from Expressive Faces, Body Gestures and Speech. In: Boukis, C., Pnevmatikakis, L., Polymenakos, L. (eds.) Artificial Intelligence and Innovations 2007: from Theory to Applications. IFIP, vol. 247, pp. 375–388. Springer, Boston (2007)

72. Batliner, A., Steidl, S., Hacker, C., Nöth, E.: Private emotions versus social interaction: A data-driven approach towards analysing emotion in speech. Journal of User Modeling and User-Adapted Interaction 18(1), 175–206 (2008)
73. Zeng, Z., Pantic, M., Roisman, G., Huang, T.: A survey of affect recognition methods: Audio, visual, and spontaneous expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(1), 39–58 (2009)
74. Glodek, M., Tschechne, S., Layher, G., Schels, M., Brosch, T., Scherer, S., Kächele, M., Schmidt, M., Neumann, H., Palm, G., Schwenker, F.: Multiple Classifier Systems for the Classification of Audio-Visual Emotional States. In: D'Mello, S., Graesser, A., Schuller, B., Martin, J.-C. (eds.) ACII 2011, Part II. LNCS, vol. 6975, pp. 359–368. Springer, Heidelberg (2011)
75. Walter, S., Scherer, S., Schels, M., Glodek, M., Hrabal, D., Schmidt, M., Böck, R., Limbrecht, K., Traue, H., Schwenker, F.: Multimodal Emotion Classification in Naturalistic User Behavior. In: Jacko, J.A. (ed.) HCI International 2011, Part III. LNCS, vol. 6763, pp. 603–611. Springer, Heidelberg (2011)
76. Schels, M., Glodek, M., Meudt, S., Schmidt, M., Hrabal, D., Böck, R., Walter, S., Schwenker, F.: Multi-modal classifier-fusion for the classification of emotional states in WOZ scenarios. In: Proceeding of the International Conference on Affective and Pleasurable Design, APD (2012)
77. Scherer, S., Glodek, M., Schwenker, F., Campbell, N., Palm, G.: Spotting laughter in natural multiparty conversations: A comparison of automatic online and offline approaches using audiovisual data. ACM Transactions on Interactive Intelligent System 2(1), 4:1–4:31 (2012)
78. Schuller, B., Valstar, M., Eyben, F., McKeown, G., Cowie, R., Pantic, M.: AVEC 2011–The First International Audio/Visual Emotion Challenge. In: D'Mello, S., Graesser, A., Schuller, B., Martin, J.-C. (eds.) ACII 2011, Part II. LNCS, vol. 6975, pp. 415–424. Springer, Heidelberg (2011)
79. Douglas-Cowie, E., Cox, C., Martin, J., Devillers, L., Cowie, R., Sneddon, I., McRorie, M., Pelachaud, C., Peters, C., Lowry, O., et al.: The HUMAINE database. In: Emotion-Oriented Systems: The Humaine Handbook, pp. 243–284. Springer (2011)
80. Frommer, J., Michaelis, B., Rösner, D., Wendemuth, A., Friesen, R., Haase, M., Kunze, M., Andrich, R., Lange, J., Panning, A., Siegert., I.: Towards emotion and affect detection in the multimodal LAST MINUTE corpus. In: Proceedings of the International Conference on Language Resources and Evaluation, LREC (2012)
81. Gnjatovic, M., Rösner, D.: On the role of the Nimitek corpus in developing an emotion adaptive spoken dialogue system. In: Proceedings of the International Conference on Language Resources and Evaluation Conference, LREC (2008)
82. Strauß, P.M., Hoffmann, H., Minker, W., Neumann, H., Palm, G., Scherer, S., Traue, H., Weidenbacher, U.: The PIT corpus of german multi-party dialogues. In: Proceedings of the International Conference on Language Resources and Evaluation, LREC (2008)
83. Douglas-Cowie, E., Cowie, R., Cox, C., Amir, N., Heylen, D.: The sensitive artificial listener: An induction technique for generating emotionally coloured conversation. In: Proceedings of the Workshop Corpora for Research on Emotion and Affect at the International Conference on Language Resources and Evaluation, LREC (2008)
84. Türk, U.: The technical processing in Smartkom data collection: A case study. In: Proceedings of the European Conference on Speech Communication and Technology, Eurospeech (2001)

85. Scherer, S., Siegert, I., Bigalke, L., Meudt, S.: Developing an expressive speech labeling tool incorporating the temporal characteristics of emotion. In: Proceedings of the International Conference on Language Resources and Evaluation (LREC). European Language Resources Association, ELRA (2010)
86. Böck, R., Siegert, I., Haase, M., Lange, J., Wendemuth, A.: ikannotate – A Tool for Labelling, Transcription, and Annotation of Emotionally Coloured Speech. In: D'Mello, S., Graesser, A., Schuller, B., Martin, J.-C. (eds.) ACII 2011, Part I. LNCS, vol. 6974, pp. 25–34. Springer, Heidelberg (2011)
87. Siegert, I., Böck, R., Philippou-Hübner, D., Vlasenko, B., Wendemuth, A.: Appropriate emotional labeling of non-acted speech using basic emotions, Geneva emotion wheel and self assessment manikins. In: Proceedings of the International Conference on Multimedia and Expo (ICME). IEEE (2011)
88. Dietrich, C., Schwenker, F., Palm, G.: Classification of time series utilizing temporal and decision fusion. Journal of Multiple Classifier Systems, 378–387 (2001)
89. Dietrich, C., Palm, G., Schwenker, F.: Decision templates for the classification of bioacoustic time series. Journal of Information Fusion 4(2), 101–109 (2003)
90. Dietrich, C., Palm, G., Riede, K., Schwenker, F.: Classification of bioacoustic time series based on the combination of global and local decisions. Journal of Pattern Recognition 37(12), 2293–2305 (2004)
91. Brown, G., Wyatt, J., Harris, R., Yao, X.: Diversity creation methods: A survey and categorisation. Journal of Information Fusion 6(1), 5–20 (2005)
92. Kahsay, L., Schwenker, F., Palm, G.: Comparison of multiclass SVM decomposition schemes for visual object recognition. Journal of Pattern Recognition, 334–341 (2005)
93. Mayer, G., Utz, H., Palm, G.: Information integration in a multi-stage object classifier. Journal of Autonome Mobile Systeme, 211–217 (2005)
94. Plumpton, C.O., Kuncheva, L.I., Oosterhof, N.N., Johnston, S.J.: Naive random subspace ensemble with linear classifiers for real-time classification of fMRI data. Journal of Pattern Recognition 45(6), 2101–2108 (2012)
95. Schwenker, F., Kestler, H., Palm, G.: Three learning phases for radial-basis-function networks. Journal of Neural Networks 14(4-5), 439–458 (2001)
96. Scherer, S., Schwenker, F., Palm, G.: Classifier fusion for emotion recognition from speech. In: Proceedings of the International Conference on Intelligent Environments (IE), pp. 152–155 (2007)
97. Schels, M., Thiel, C., Schwenker, F., Palm, G.: Classifier Fusion Applied to Facial Expression Recognition: An Experimental Comparison. In: Ritter, H., Sagerer, G., Dillmann, R., Buss, M. (eds.) Human Centered Robot Systems. Cognitive Systems Monographs, vol. 6, pp. 121–129. Springer, Heidelberg (2009)
98. Scherer, S., Schwenker, F., Palm, G.: Classifier fusion for emotion recognition from speech. In: Proceeding of the International Conference Intelligent Environments (IE), pp. 95–117. Springer (2009)
99. Scherer, S., Trentin, E., Schwenker, F., Palm, G.: Approaching emotion in human computer interaction. In: Proceedings of the International Workshop on Spoken Dialogue Systems (IWSDS), pp. 156–168 (2009)
100. Scherer, S., Schwenker, F., Campbell, N., Palm, G.: Multimodal Laughter Detection in Natural Discourses. In: Ritter, H., Sagerer, G., Dillmann, R., Buss, M. (eds.) Human Centered Robot Systems. Cognitive Systems Monographs, vol. 6, pp. 111–120. Springer, Heidelberg (2009)
101. Schuller, B., Vlasenko, B., Eyben, F., Wöllmer, M., Stuhlsatz, A., Wendemuth, A., Rigoll, G.: Cross-corpus acoustic emotion recognition: Variances and strategies. IEEE Transactions on Affective Computing 1(2), 119–131 (2010)

102. Wöllmer, M., Metallinou, A., Eyben, F., Schuller, B., Narayanan, S.: Context-sensitive multimodal emotion recognition from speech and facial expression using bidirectional lstm modeling. In: Proceedings of the Annual Conference of the International Speech Communication Association (ISCA), Interspeech, pp. 2362–2365 (2010)
103. Esparza, J., Scherer, S., Schwenker, F.: Studying Self- and Active-Training Methods for Multi-feature Set Emotion Recognition. In: Schwenker, F., Trentin, E. (eds.) PSL 2011. LNCS, vol. 7081, pp. 19–31. Springer, Heidelberg (2012)
104. Blum, A., Mitchell, T.: Combining labeled and unlabeled data with co-training. In: Proceedings of the Annual Conference on Computational Learning Theory (COLT), pp. 92–100. Morgan Kaufmann (1998)
105. Bennet, K., Demiriz, A., Maclin, R.: Exploiting unlabeled data in ensemble methods. In: Proceedings of the International Conference on Knowledge Discovery and Data Mining (KDD), pp. 289–296 (2002)
106. Grandvalet, Y., Bengio, Y.: Semi-supervised learning by entropy minimization. Journal of Advances in Neural Information Processing Systems 17, 529–536 (2005)
107. Lawrence, N., Jordan, M.: Semi-supervised learning via Gaussian processes. Journal of Advances in Neural Information Processing Systems 17, 753–760 (2005)
108. Schwenker, F., Dietrich, C., Thiel, C., Palm, G.: Learning of decision fusion mappings for pattern recognition. Proceedings of the International Journal on Artificial Intelligence and Machine Learning (AIML) 6, 17–21 (2006)
109. Thiel, C., Scherer, S., Schwenker, F.: Fuzzy-Input Fuzzy-Output One-Against-All Support Vector Machines. In: Apolloni, B., Howlett, R.J., Jain, L. (eds.) KES 2007, Part III. LNCS (LNAI), vol. 4694, pp. 156–165. Springer, Heidelberg (2007)
110. Wang, W., Zhou, Z.-H.: Analyzing Co-training Style Algorithms. In: Kok, J.N., Koronacki, J., Lopez de Mantaras, R., Matwin, S., Mladenič, D., Skowron, A. (eds.) ECML 2007. LNCS (LNAI), vol. 4701, pp. 454–465. Springer, Heidelberg (2007)
111. Abdel Hady, M.F., Schwenker, F.: Decision Templates Based RBF Network for Tree-Structured Multiple Classifier Fusion. In: Benediktsson, J.A., Kittler, J., Roli, F. (eds.) MCS 2009. LNCS, vol. 5519, pp. 92–101. Springer, Heidelberg (2009)
112. Abdel Hady, M.F., Schels, M., Schwenker, F., Palm, G.: Semi-supervised Facial Expressions Annotation Using Co-Training with Fast Probabilistic Tri-Class SVMs. In: Diamantaras, K., Duch, W., Iliadis, L.S. (eds.) ICANN 2010, Part II. LNCS, vol. 6353, pp. 70–75. Springer, Heidelberg (2010)
113. Abdel Hady, M.F., Schwenker, F., Palm, G.: Semi-supervised learning for tree-structured ensembles of RBF networks with co-training. Neural Networks 23(4), 497–509 (2010)
114. Abdel Hady, M.F., Schwenker, F., Palm, G.: When classifier selection meets information theory: A unifying view. In: Proceedings of the International Conference on Soft Computing and Pattern Recognition (SoCPaR), pp. 314–319. IEEE (2010)
115. Abdel Hady, M., Schwenker, F.: Combining committee-based semi-supervised learning and active learning. Journal of Computer Science and Technology (JCST): Special Issue on Advances in Machine Learning and Applications 25(4), 681–698 (2010)
116. Abdel Hady, M., Schwenker, F., Palm, G.: Semi-supervised learning for tree-structured ensembles of RBF networks with co-training. Journal of Neural Networks 23(4), 497–509 (2010)

117. Zhu, X.: Semi-supervised learning literature survey. Technical Report 1530, Department of Computer Sciences, University of Wisconsin at Madison (2008)
118. Settles, B.: Active learning literature survey. Technical Report 1648, Department of Computer Sciences, University of Wisconsin-Madison (2009)
119. El Gayar, N., Schwenker, F., Palm, G.: A study of the robustness of KNN classifiers trained using soft labels. Journal of Artificial Neural Networks in Pattern Recognition, 67–80 (2006)
120. Thiel, C., Sonntag, B., Schwenker, F.: Experiments with Supervised Fuzzy LVQ. In: Prevost, L., Marinai, S., Schwenker, F. (eds.) ANNPR 2008. LNCS (LNAI), vol. 5064, pp. 125–132. Springer, Heidelberg (2008)
121. Thiel, C., Giacco, F., Schwenker, F., Palm, G.: Comparison of neural classification algorithms applied to land cover mapping. In: Proceeding of the International Conference on New Directions in Neural Networks: Italian Workshop on Neural Networks: (WIRN), pp. 254–263. IOS Press (2009)