Proceedings of the 2014 IEEE
International Conference on Robotics and Biomimetics
December 5-10, 2014, Bali, Indonesia

Multi-modal Biometrics for Real-Life Person-Specific Emotional Human-Robot-Interaction

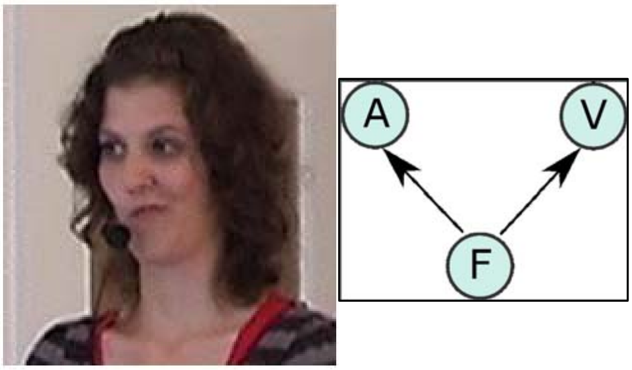
Ahmad Rabie and Uwe Handmann

***Abstract*—Currently in home environments, robot assisting systems with emotion understanding ability are generally achieved in two several manners. The first is the implementing of such systems in such a way that they offer general services for all considered persons without considering privacy, special needs of their interaction partners. The second way is the targetting of such systems for merely one person. In this work we present a robot assisting system, which has both the abilities of assisting several persons at the same time and sustaining their privacy and security issues. The robot can interact with its interaction partner emotionally by analyzing the emotions of her expressed either visually, facial expression, or auditive, speech prosody. The role of this system is the providing of person-specific support in home environment. In order to identify its interaction partner the system uses diverse biometric traits. According to the recognized ID the system, first, adopts towards the needs of recognized person. Second the system loads the corresponding emotional profile of the detected interaction partner in order to practice a person-specific emotional human-robot interaction, which has an advantage over the person independent interaction.**

I. Introduction

RECOGNIZING emotions is widely accepted as one relevant step towards more natural interaction in human-robot- and, more general, human-machine-interaction. The new scientific understanding of emotions on the one hand, and the rapid evolution of computing system skills on the other, provided inspiration to numerous researchers to build machines that will have the ability to recognize, express, model, and communicate emotions. Integrating such robot systems in modern smart-home environment gains the interest of numerous researchers. The use of these systems ranges from supporting by ordinary daily life activities, such as preparing a cup of tea, to serving as a wizard by self-learning software programs to healthy supporting of elderly and disabled people [1]. Fig 1(a) presents an example from a human-robot interaction study in life-like scenario.

However, systems with emotion understanding ability perform currently in two basic modes. The person-specific methods employs a pre-step of identifying the interaction partner and according to the recognized ID the corresponding emotional profile of her. The person independent one uses a general emotional profile for all possible interaction partners. While the person-specific method outperforms the person independent method, the latter exhibits more validity than the former [2], [3].



(a) (b)

Fig 1: (a) Emotion articulation in an HRI task. (b) The Bayesian network structure for decision-level fusion.

This paper first briefly introduces both robot platform and a biometric system integrated in the home environment will be emerged in order to build a robust person-specific emotion analysis system for a robot in home environment (Sec. II). The relevance and related work on multi-modal person recognition systems in real-life home environments will be discussed in Sec. III. Afterworlds, our systems for visual and acoustic emotion recognition are introduced in Sections IV.A and IV.B respectively. A probabilistic fusion method is used to integrate both the systems in a bimodal emotion recognition system that recognizes the six basic Ekmanian emotions (anger, disgust, happiness, fear, sadness, and surprise) plus the neutral emotion state, see Sec. IV. The pre-step of identifying of the robot's interaction partner, which achieved using a multi-modal biometric system, as well as the integration of the robot within the whole are presented in Sec.V Discussion and future work will conclude this work.

II. HARDWARE

The robot platform is equipped with one visual and one audio system. The visual system, two pan-tilt cameras mounted on the head of the robot, is used to detect, locate and identify the interaction partner and her emotional states

Manuscript received August 8, 2014. This work was supported in part by the Ministerium für Innovation, Wissenschaft und Forschung des Landes NRW, Germany

Ahmad Rabie is with the Computer Sciences Institute, University of Applied Sciences Ruhr-West, Bottrop, Germany, (e-mail: ahmad.rabie@hs-ruhr-west.de). Uwe Handmann is with the Computer Sciences Institute, University of Applied Sciences Ruhr-West, Bottrop, Germany (e-mail: uwe.handmann@hs-ruhr-west.de).

through the analyzing her face images. The audio system does the same role excepting the identification but by employing the audio signal of speech instead. Utilizing both the modalities will give for the robot the ability of multi-modal analyzing the emotional state of her interaction partner expressed via accompanying facial expressions or speech prosody. The basic goal presented in this paper is the integration and optimization of a person recognition module that enables for the robot the identification of her interaction partner as a pre-step of person-specific emotional human-robot-interaction. The robot's interaction partner will be additionally detected and identified by further biometric subsystems integrated in the surrounding environment, such as iris-, hand-palm-vein and finger-vein-based identification systems. These biometric modalities will be fused with the ones of face images and speech signals will enhance the performance of the whole system in the case of noisy input data or the absence of the considering modality.

III. MULTI-MODAL BIOMETRIC SYSTEM

Uni-modal biometric systems is challenged by multiple issues, such as noisy captured data, non-universality, upper bound of identification accuracy, man-in-middle attacks and the absence of the considered modality. Some of these limitations of uni-modal systems can be avoided by realizing multi-modal system, which fuses input data of multiple biometric traits. This can be accomplished by fusing two (bimodal) or more biometric traits (multi-modal) at several levels. In signal fusion level data from multiple sources are fused, as example raw data obtained using multiple sensors or multiple snapshots using single sensor. Multiple feature sets, which originate from multiple feature extracting algorithms, are gathered, normalized, transformed and dimensionally reduced to build a single feature set in the feature level fusion mode. In score fusion level multiple scores of multiple matchers will be combined in order to get a final score about the similarity or the distance of the captured identity to those already saved in data base. Final decisions of multiple systems can be logically fused in the decision fusion level in order to get the final decision of the multi-modal system. As the aim of our work is the offering for a robot ability of robust identifying of her interaction partner in the case of noisy input data, a score level fusion system is used. The fusion method performs in two several modes. A Bayesian-based probabilistic score-level-fusion method, as illustrated in Fig 2, is used to fuse the scores of all available traits in the case of being more than one biometric trait available. In the case of being only one biometric trait available it suffices when the person is identified using it. That means, logically, that the decisions of applied stand-alone biometric systems (uni-modals) have to be joined using simple OR rule. Therefore, we have implemented our multi-modal system in such away, that it gives the ID of the interaction partner as soon as one or more biometric treats are recognized. Fig 2(b) illustrates the simple logical OR fusion method used to integrate multiple biometric subsystems in a single multimodal one.

From the well-known biometric treats we used face, finger vein and hand palm vein. The reason for this is that these features don't demand any direct contact to the used sensor, which serves our goal of having a touch-less assisting system. The following subsections give small explanation of the used uni-modal system and a detailed explanation of the fusion method.

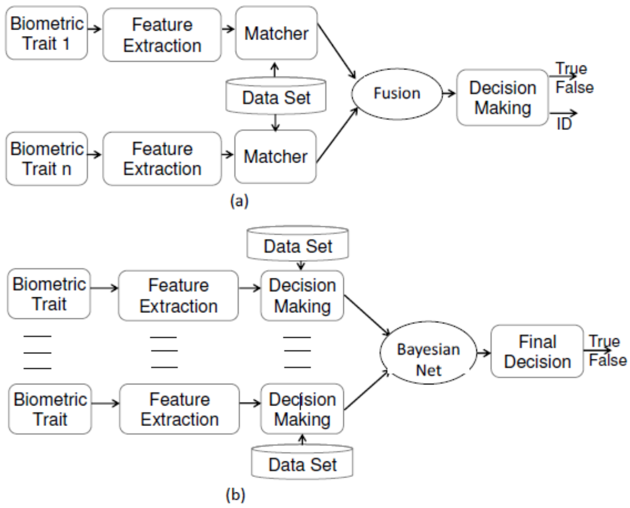


Fig 2: a) Score level fusion method used to combine multiple biometric Traits. (b) Simple Bayesian net used to combine multiple uni-modal biometric systems in a single multi-modal one

A. *Face Recognition*

Face recognition is one of the most populated and almost the most researched method for person authentication. Not few face recognition systems have been developed for automatically recognizing faces from either still or video images [4] [5]. For our system a robust, full automatic and real-life face-recognition-based person recognizer is employed [6]. The basic technique applied here are Active Appearance models (AAMs) First introduced by Cootes *et. al.* [7]. The generative AAM approach uses statistical models of shape and texture to describe and synthesize face images. An AAM, which is built from training set, can describe and generate both shape and texture using a single appearance parameter vector, which is used as feature vector for the classification. The "active" component of an AAM is a search algorithm that computes the appearance parameter vector for a yet unseen face iteratively, starting from an initial estimation of its shape. The AAM fitting algorithm is part of the integrated vision system [6] that consists of three basic components. Face pose and basic facial features (BFFs), such as nose, mouth and eyes, are recognized by the face detection module [8]. This face detection in particular allows applying the AAM approach in real-world environments as it has proven to be robust enough for face identification in ordinary home environments [5]. The coordinates representing these features are conveyed to the facial feature extraction module. Here, the BFFs are used to initialize the iterative AAM fitting algorithm. After the features are extracted the resulting parameter vector for every image frame is passed to a classifier which perform in either identification mode, comparing the extracted feature vector with feature vectors of all already saved identities, or verification mode, comparing the extracted feature vector according to a claimed identity. Besides the feature vector, AAM fitting also returns a reconstruction error that is applied as a confidence measure to reason about the quality of the fitting and also to reject prior false positives resulting from face detection. A one-against-all Support Vector Machine [9] is applied as classifier.

B. Hand Palm Vein Recognition

Typically, palm vein recognition system performs three basic tasks, namely image acquisition, feature extraction and decision making. Image preprocessing and image enhancements could be achieved in order to get features with reliable quality for the next step of classification. For on-line capturing of palm vein images an M2Sys scanner is used. This device uses a near infrared light to create a vein-map of the users' palm, which serves as a biometric feature. It scans arteries beneath the skin. Therefore it is practically impossible these templates to be forged through creating someone else's biometric template. The device works in a contact less mode, in which the user has not to touch the sensor directly. For the stages of feature extracting and matching an algorithm similar to the one presented by [10] is utilized. Extracting the region of interest (ROI) from the captured palm vein image is an essential step of preprocessing. For this goal the inscribed circle-based segmentation which extracts the ROI from the original palm vein image is used. The basic step toward getting that is to calculate the inscribed circle that meets the boundary of a palm so that it can extract as large an area as possible from the central part of the palm vein image. First, an edge detecting method is used to obtain the contour of the palm. Using the contour of the palm the biggest inscribed circle is then calculated. Once the circle is determined, ROI image is smoothed by using the standard deviation Gaussian kernel filter. In order to reduce some high frequency noise, ROI image is then smoothed by the Gaussian smooth filter. Local contrast enhancement is then applied in order to blurred ROI image caused by Gaussian filtering. For the extraction of vein-pattern-based features (vein length and minutiae) from preprocessed images a minutiae extracting method, which is basically employed in finger print recognition systems, is adopted. This method performs in four sub-steps. First, binarization is achieved using the local threshold scheme. A median filter is then used to reduce the noise. Finally, the morphological thinning method is used to thin and repair the vein line and the position information of the minutiae can be got. A minutiae based matching method, in which the position and the orientation of each corresponding couple of minutiae are compared, is based for decision making [11].

C. Finger Vein Recognition

Like palm vein recognition system finger vein recognition systems consist of three basic components, namely image capturing, feature extraction and decision making. A suitable scanner, which employed infra-red technology from Hitachi, is used. This scanner captures image of the vein inside the finger, therefor the captured images are virtually impossible to replicate. The scanner works by passing near-infrared light through the finger. This is partially absorbed by the hemoglobin in the veins, allowing an image to be recorded on a CCD camera. Unless the location and the orientation of the finger within the capturing device is explicitly predefined a step of image normalization has to be conducted. Acting on the assumption that the veins in a finger vein image could be seen as lines with higher gray values as the rest of the image, the task of detecting such vein could be seen as a task of following lines within image. Line tracking offer us the ability of doing that robustly [12]. The line-tracking process starts at any pixel in the captured image. The current pixel position is called the ``current growth point", which moves pixel by pixel along the dark line. The direction of movement depends on the results of checking gray values of the surrounding neighborhood.

The lowest gray value of the cross-sectional profile, which represents the depth of the profile, is checked around the current tracking point. If pixel p is a neighbor of the current tracking point and the cross-sectional profile on this pixel looks like a valley bottom, then the current tracking point is considered to be located on a dark line. The angle between horizontal line and the line that connecting the current growth point and the considered neighboring pixel is called Θ. In order to detect the direction of the dark line the depth of the valley is scanned with varying angle Θ_i. The highest value of defines the direction of the dark line. After that, the current growth point moves to the closest pixel toward this direction and the process is repeated iteratively. In the case of no detecting the valley in any direction Θ_i then current growth point does not belong to any dark line and the tracking operation starts considering another position as current growth point. Toward detecting multiple veins in the image multiple vein tracking sequences start at various positions simultaneously. The results of tracking are stored in a matrix of the same size of the original image, which is called "locus space". Each entry of the matrix contains information about how much the corresponding pixel of the original image is tracked. Entries of matrix with high values mean that the corresponding pixels of the original image have high probability of being belonging to a vein. The matrix is then binarized by utilizing a thresholding technique. Spatial reduction and relabeling are then applied on the binarized image in order to retain the vein line width as small as about 3 pixels in the image. Finally, a conventional template matching technique is applied to get the final decision about the between the captured vein data and the already registered one [12].

IV. Bi-Modal Emotion Recognition

Theories of modality fusion in human perception do not agree on how information from different modalities should be integrated. For example, the Fuzzy Logical Model of Perception (FLMP) [13] states that stimuli from different modalities should be treated as independent sources of information and be combined regardless of the kind of information they contain. This view is not undisputed (i.e. [14]) and it has been argued that the FLMP does not work well when confronted with conflicting information from different modalities [15]. Perceptual results suggest that, at least for the case of emotion recognition, the modalities should be weighted according to which information that they convey best [16]: the visual modality primarily transmits valence (positive or negative value) whereas the auditory channel mainly contains information about activation.

In our work we challenge this approach by analyzing the auditory and visual stimuli with respect to their general discriminative power in recognizing emotions. Note that in our work we focus on interactive scenarios and are thus targeting at systems that are able to work on-line. The approaches we present in this paper are, therefore, not only being tested off-line on existing databases but have proven their applicability in robotic applications in real world settings [17], [18]. This is in contrast to other work (e.g. [19], [20]), which has focused on off-line emotion recognition only. The following three sections will provide a brief introduction on the respective uni-modal analysis techniques as well on the proposed probabilistic decision level fusion.

A. *Visual Facial Expression Recognition*

In order to recognize basic emotion visually, we take a closer look into the interlocutor's face. The basic technique applied here is an extension of Active Appearance models (AAMs) applied for biometric face recognition, see Sec. III.A. All steps are the same up to getting the feature vector extracted. After the features are extracted the resulting parameter vector for every image frame is passed to a classifier which categorizes it in one of the six basic emotions in addition to the neutral one. Besides the feature vector, AAM fitting also returns a reconstruction error that is applied as a confidence measure to reason about the quality of the fitting and also to reject prior false positives resulting from face detection. As classifier a one-against-all Support Vector Machine is applied. The whole system is applicable in soft real-time, running at a rate of approximate (5) Hz on recent PC hardware.

B. *Emotion Recognition From Speech*

For the recognition of emotions from speech, EmoVoice, a framework that features off-line analysis of available emotional speech databases, as well as on-line analysis of emotional speech for applications, is used [21]. The approach taken there is purely based on acoustic features, that is no word information is used. As a first step in feature extraction, a large vector of statistical features based on prosodic and acoustic properties of the speech signal was calculated for each utterance in the DaFEx database [22]. From this large vector of over 1400 features the most relevant ones were selected by correlation-based feature subset selection [23]. This selection is necessary to increase performance as well as speed of classification. By this way, 71 features related to pitch, energy, MFCCs, to linear regression and range of the frequency spectrum of short-term signal segments, to the speech proportion and to the length of voiced and unvoiced parts in an utterance, and the number of glottal pulses remained. The full procedure of extracting features is described in [24], [21]. For classification, again support vector machines were used, but with a linear kernel. The feature selection is typically done off-line, but the feature extraction and classification can be done in real-time. Utterances as classification units, which are normally not available in on-line applications, can be replaced by an on-the-fly segmentation into parts with voice activity.

C. *Probabilistic Decision Level Fusion*

As affective states in interaction are usually conveyed on different cues at the same time, we agree with other works summarized in [25] that a fusion of visual and acoustic recognition yields significant performance gains. Hence, we followed the idea of an on-line integration scheme based on the prior off-line analysis of recognition results on a database. In current classification fusion research, usually two types of multi-modal fusion strategies are applied namely feature level fusion and decision level fusion. Both types combine different modalities of data to achieve better recognition performance. In the former one, the feature spaces of all modalities are merged into one feature space, which is then conveyed to a single classifier. While in the latter type the classification is performed on each modality separately, then the results of each modality are fused to final class-prediction accuracy. Due to the inherently different nature of our visual and acoustic cues, we decided for a decision-level fusion scheme. But instead of applying majority voting or other simple fusion techniques, we explicitly take the performance of each individual classifier into account and weight it according to their respective discrimination power.

The proposed probabilistic approach for this fusion are Bayesian networks with a rather simple structure depicted in Fig 1(b). Based on the classification results of the individual visual and acoustic classifiers, we feed these into the Bayesian network as *evidences* of the observable nodes (A and V, respectively). By Bayesian inference the posteriori probabilities of the unobservable affective fusion (F) node are computed and taken as final result.

The required probability tables of the Bayesian network are obtained from a performance evaluation of the individual classifiers in an off-line training phase based on ground-truth annotated databases. Therefore, confusion matrices of each classifier are turned into probability tables modeling and communicates with the interaction partner emotionally. Face, finger vein or hand palm vein data are acquired from the considered interaction partner using a camera fixed behind the mirror or the door of the first aid box in bathroom or a finger vein sensor encapsulated within blood

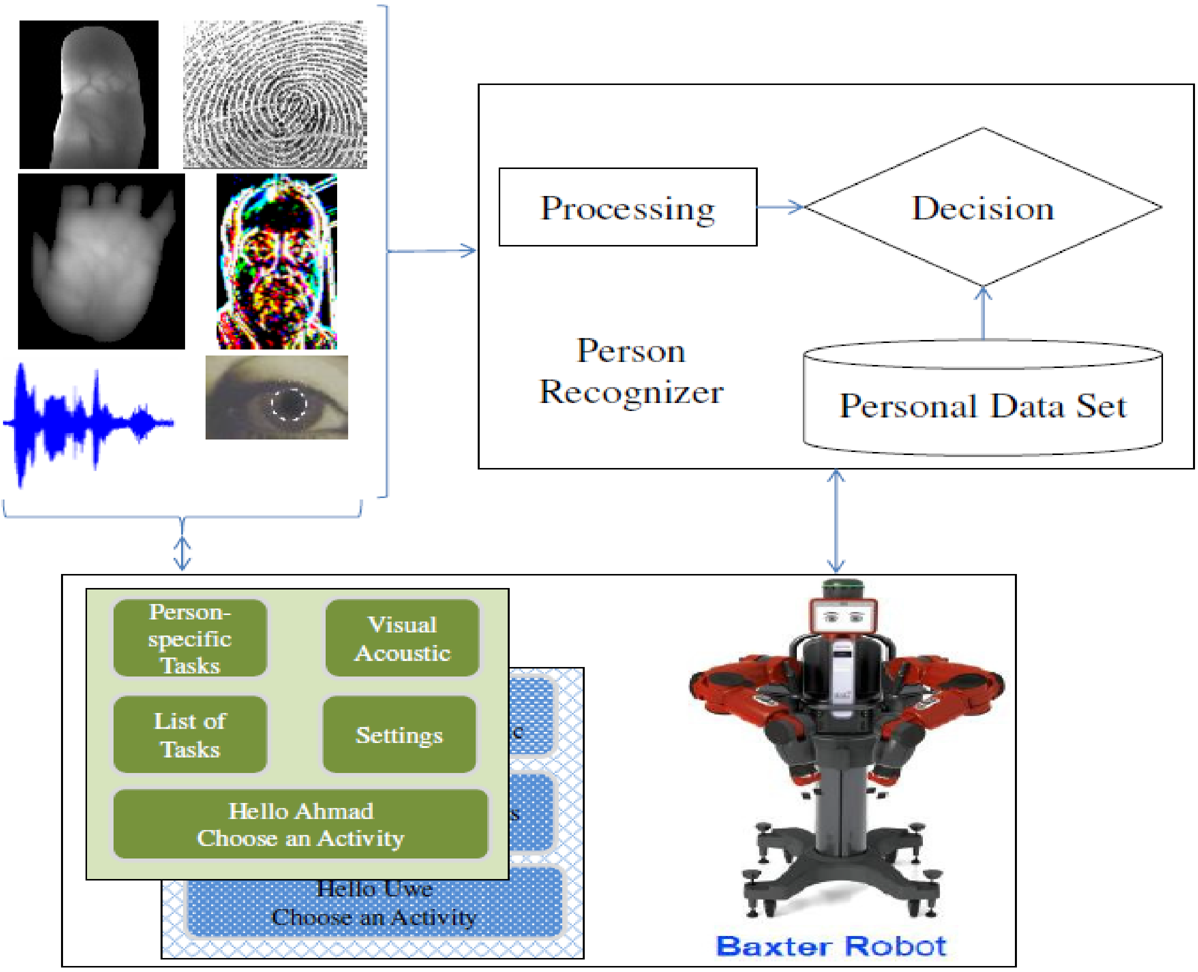


Fig 3: Basic architecture of whole system. Person is identified by person recognizer inserted in the surrounding environment or in robot platform. Robot interacts with its interaction partner according to the recognized ID.

the dependent observation probabilities of the model according to the arrows in Fig 1(b). In the notion of Zeng *et al.* [25], our fusion scheme is referred to as model-level instead of decision-level fusion, as it takes the respective classification performance models into account.

V. SYSTEM INTEGRATION

The basic structure of the whole system could be divided in two basic subsystems, namely person recognition subsystem implemented in the surrounding Environment, i.e. household equipment, and a robot, as illustrated in Fig 3. Person recognition system integrated in the surrounding environment is based on the analysis of one or more biometric treats in order to identify/verify the interaction partner. The robot system has the two roles of identifying the interaction partner via face images, if available, support glucose meter or by a hand palm vein sensor hidden in a hand air dryer [25]. Face images could be captured from the robot's cameras too. Once one or more of the above mentioned biometric features of the interaction partner are captured, either deliberately or accidentally, the suitable biometric features are extracted and analyzed by the corresponding person recognition subsystem and the final decision of the system is delivered as a person- ID to the medical organizer. After the interaction partner is identified the robot modifies first the theme of the interface in such a way, that it fulfills the needs of the recognized person. A list of person-specific tasks will be presented either visually (on a screen) or acoustically (through speaker). Via the established interface the robot will request tasks from the interaction partner. The robot will then try to implement the requested task and use the experienced emotion state of the interaction partner as a scale of satisfaction of her.

In practice, the biometric subsystem is implemented on a normal PC equipped with suitable biometric devices, as discussed in Sec. III. The communication between the biometric subsystem and the robot is achieved via TCP-connection. A synchronization process is iteratively done in order to get both person data base, which is saved on the normal PC, and a emotional profiles, which are saved on the robot platform, synchronized permanently. When the interaction partner is identified by the biometric system the ID of this person will be sent to the robot in order to fetch the corresponding emotional profile.

VI. CONCLUSION AND OUTLOOK

In this paper we presented our approach of integrating a person identification system and a robot in a person-specific robot assisting system for home environment. As we strive to provide for people sharing a single household a touch-free medical assistant system, we presented a robot assisting system that fulfills the requirements and offering for them a reliable alleviation their ordinary life situations. The robot has the ability of interacting with its interaction partner emotionally. Person-specific emotional profile and task list are loaded according to the person ID provided by either the biometric subsystem, person recognition system of the robot via face images or both. An open issue of the system is giving the ability for the robot of offering healthy support for its interaction partners. To solve this problem we focus for the next step on the implementing of person-specific medical recognition system on the robot [27], [1]. The bimodal emotion analysis system perform upon the assumption that both modalities, facial expression and speech prosody, are available during the interaction. A future addition to the system could be an extension of the fusion method in such a way, that it detects if both modalities are available or not and adapts accordingly.

ACKNOWLEDGMENT

This work was partly funded by the Ministerium für Innovation, Wissenschaft und Forschung des Landes NRW, Germany

REFERENCES

[1] Rabie, A. and Handmann, U. *NFC-Based person-specific medical assistant for elderly and disabled people*. In proceedings of International Conference on Next Generation Computing and Communication Technologies, 2014, pp. 238-243.

[2] A. Rabie, T. Vogt, M. Hanheide, and B. Wrede, "*Evaluation and discussion of multi-modal emotion recognition*", in ICCEE, 2009.

[3] A. Rabie and U. Handmann, "*Fusion of audio- and visual cues for real-life emotional human robot interaction",* in Annual Conf. of the German Association for Pattern Recognition, 2011.

[4] U. Handmann, S. Hommel, M. Brauckmann, and M. Dose, Towards Service Robots for Everyday Environments, ch. *Face Detection and Person Identification on Mobile Platforms*, pp. 227–234. Springer, 2012.

[5] M. Hanheide, S. Wrede, C. Lang, and G. Sagerer, "Who am i talking with? a face memory for social robots," 2008.

[6] A. Rabie, C. Lang, M. Hanheide, M. Castrillon-Santana, and G. Sagerer, "*Automatic initialization for facial analysis in interactive robotics*", in ICVS, 2008.

[7] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "*Active appearance models*", PAMI, vol. 23, pp. 681–685, June 2001.

[8] M. Castrillón, O. Déniz, C. Guerra, and M. Hernández, "*Encara2: Real-time detection of multiple faces at different resolutions in video streams*", Journal of Visual Communication and Image Representation, vol. 18, no. 2, pp. 130–140, 2007.

[9] B. Schölkopf and A. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. Cambridge Press, 2002.

[10] H. Zhang and D. Hu, "*A palm vein recognition system*", in International Conference on Intelligent Computation Technology and Automation, 2010.

[11] X.-F. Tong, P.-F. Li, and Y.-X. Han, "*An algorithm for fingerprint basic minutiea feature matching*", in International Conference on Machine Learning and Cybernetics, 2012.

[12] N. Miura, A. Nagasaka, and T. Miyatake, "*Feature extraction of fingervein patterns based on repeated line tracking and its application to personal identification*", Machine Vision and Applications archive, vol. 15, pp. 194–203, 2004.

[13] D. W. Massaro and P. B. Egan, "*Perceiving affect from the voice and the face*", Psychonomoic Bulletin and Review, vol. 3, pp. 215–221.

[14] B. de Gelder and J. Vroomen, "*Bimodal emotion perception: integration across separate modalities, cross-modal perceptula grouping or perception of multimodal events?*", Cognition and Emotion, vol. 14, no. 3, pp. 321–324, 2000.

[15] J. L. Schwartz, "*Why the FLMP should not be applied to McGurk data .. or how to better compare models in the bazesian framework*", in Proc. Int. Conf. Audio-Visual Speech Processing, pp. 77–82, 2003.

[16] S. Fagel, "*Emotional mcgurk effect*", in Proc. Int. Conf. on Speech Prosody, (Dresden, Germany), 2006.

[17] A. Rabie, C. Lang, M. Hanheide, M. Castrillon-Santana, and G. Sagerer, "Automatic initialization for facial analysis in interactive robotics," 2008.

[18] F. Hegel, T. Spexard, T. Vogt, G. Horstmann, and B. Wrede, "*Playing a different imitation game: Interaction with an empathic android robot*", in Proc. Int. Conf. Humanoid Robots, pp. 56–61, 2006.

[19] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "*Analysis of emotion recognition using facial expressions, speech and multimodal information*", in ICMI, 2004.

[20] G. Caridakis, L. Malatesta, L. Kessous, N. Amir, A. Raouzaiou, and K. Karpouzis, "*Modeling naturalistic affective states via facial and vocal expressions recognition*", in Proc. Int. Conf. Multimodal Interfaces, (New York, NY, USA), pp. 146–154, ACM, 2006.

[21] T. Vogt, E. Andr´e, and N. Bee, "*Emovoice — A framework for online recognition of emotions from voice*", in Proc. Workshop on Perception and Interactive Technologies for Speech-Based Systems, (Irsee, Germany), June 2008.

[22] A. Battocchi, F. Pianesi, and D. Goren-Bar, "*A first evaluation study of a database of kinetic facial expressions (dafex)*", in Proc. Int. Conf. Multimodal Interfaces, pp. 214–221, ACM Press, 2005.

[23] M. A. Hall, "Correlation-based feature subset selection for machine learning," Master's thesis, University of Waikato, New Zealand, 1998.

[24] T. Vogt and E. Andr´e, "*Comparing feature sets for acted and spontaneous speech in view of automatic emotion recognition*", in Proc. of IEEE Int. Conf. on Multimedia & Expo, (Amsterdam, The Netherlands), July 2005.

[25] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "*A survey of affect recognition methods: Audio, visual, and spontaneous expressions*", IEEE Transaction on Pattern Analysis and Macine Intellegence, vol. 31, pp. 39–58, 2009.

[26] "Fraunhofer inhaus zentrum www.inhaus.fraunhofer.de/de/downloads/broschuere.html."

[27] A. Rabie and U. Handmann, "*Nfc-based person-specific assisting system in home environment*", in Proceedings of 11th World Congress on Intelligent Control and Automation, 2014.