Privacy Aware Person-specific Assisting System for Home Environment

Ahmad Rabie and Uwe Handmann
Institute of Computer Sciences, University of Applied Sciences Ruhr-West
Bottrop, Germany
{*ahmad.rabie, uwe.handmann*}@*hs-ruhrwest.de*

Keywords: Privacy and Security, Multimodal Biometrics, Assisting Systems, Data Fusion.

Abstract: As smart homes are being more and more popular, the needs of finding assisting systems which interface between users and home environments are growing. Furthermore, for people living in such homes, elderly and disabled people in particular and others in general, it is totally important to develop devices, which can support and aid them in their ordinary daily life. We focused in this work on sustaining privacy issues of the user during a real interaction with the surrounding home environment. A smart person-specific assistant system for services in home environment is proposed. The role of this system is the assisting of persons by controlling home activities and guiding the adaption of Smart-Home-Human interface towards the needs of the considered person. At the same time the system sustains privacy issues of it's interaction partner. As a special case of medical assisting the system is so implemented, that it provides for elderly or disabled people person-specific medical assistance. The system has the ability of identifying its interaction partner using some biometric features. According to the recognized ID the system, first, adopts towards the needs of recognized person. Second the system represents person-specific list of medicines either visually or auditive. And third the system gives an alarm in the case of taking medicament either later or earlier as normal taking time.

1 INTRODUCTION

Currently, assisting of people in home environment are generally achieved by either employing assisting systems which offer general services for all considered persons without considering their privacy and special needs (R.A.Ramlee et al., 2013; Dohr et al., 2010) or by using systems targeted for lonely one person (Ouchi et al., 2004). In this work we present a person-specific assisting system, which is aimed to assist several persons and sustain their privacy and security issues at the same time. In order for an assistant system to be having the ability of providing person-dependent services a prior step of identifying the interaction partner has to be achieved. Person identification/verification will be achieved using special biometric systems equipped with proper biometric sensors. Th ID of this person will then be delivered to the actual assisting system, which could be any smart device, such as smart mobile phone or PC-tablet.

1.1 Biometrics

Wide range of biometric applications are currently available on the market, under which surveillance systems, cash terminals with biometrics analysis abilities, biometrics-based payment systems (Yang, 2010), accessing digital systems, such as PCs, mobile phones and cars, accessing online services, such as online banking and person-specific services in home environment (Rabie and Handmann, 2014a; Rabie and Handmann, 2014b).
Biometrics are used either in a stand-alone mode (unimodal) or in a fused mode (multi-modal). Most commonly used biometric traits are face, fingerprint, finger vein, hand palm, iris/retina, and voice. Selecting the proper biometric trait depends on the aimed application and environment. For instance, adequate fingerprint samples require user cooperation; whereas, the face and iris images can be captured occasionally by a surveillance camera. As our goal is the building an assisting system for ordinary daily life in home environment we decided the using of traits of face, finger vein and hand palm vein, which fulfill the the requirements of being efficient, touch less and to be captured accidentally.

1.2 Multimodal Biometrics

Multimodal information fusion is the task of combining some interrelated information from multiple modalities.Fusion of multiple modalities can improve

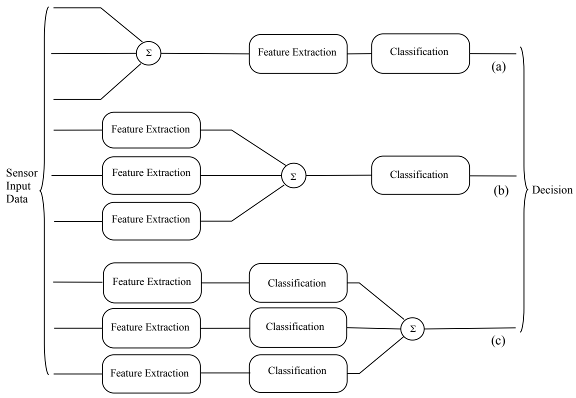


Figure 1: Three basic fusion methods used in the current multimodal emotion recognition systems. $\sum$ presents signal, features and decision fusion levels in a, b and c respectively.

the performance of the multi-modal system comparing to uni-modal systems, whose performance could be degraded by noise or illumination. It can reduce the number of false matches, which are caused by non-robust stand-alone biometric systems as well. In a person identification system, while a uni-modal system incorporates features of a single modality (face, audio, finger print, iris, ...) the multi-modal systems use information from multiple different modalities simultaneously. In current fusion research, three types of multi-modal fusion strategies are usually applied, namely data-/signal-level fusion, feature-level fusion, and decision-level fusion. Fig 5.4 depicts the three possible levels of multimodal information fusion. Signal-level fusion is applicable solely to sources of the same nature and tightly synchronous. Generally it is achieved by mixing two or more physical signals of the same nature (two auditive signals, two visual signals of two cams, two camera snapshots, etc). This type of mixing is not feasible for multimodal fusion due to the fact that different modalities always have different captors and different signal characteristics (auditive and visual). Feature-level fusion means concatenation of the features outputted from different signal processors together to construct a combined feature vector, which is then conveyed to the decision maker. It is used when there is evidence of class-dependent correlation between the features of multiple sources. For example, features can be extracted from a video processor and speech signal. Feature-level fusion is criticized for ignoring the differences in temporal structure, scale and metrics. Although, feature-level fusion demands synchronization of some extent between modalities. Another drawback of such a fusion strategy is that it is more difficult and computationally more intense than combining at the decision level. This is because of the increasing feature vector dimension, which consequently influences the performance of the whole system negatively when real time applications are aimed to. The third fusion strategy combines the semantic information captured from the individual uni-modal systems, rather than mixing together features or signals. Due to the advantages of (I) being free of synchronization issues between modalities, (II) using relative simple fusion algorithms, (III) their low computational requirement in contrast to the feature-based methods and (IV) being able to join multiple traits of different nature decision-level fusion methods are widely adopted in the field of multi-modality person recognition (Ross and Jain, 2003).The basic structure of each used modality as well as the used method of fusing multiple uni-modal biometric subsystems in a mere multimodal one will be discussed in chapter 2. Person-specific service provider will be presented in chapter 3. NFC-based medical organizer as a special case study will be discussed in chapter 3. Outlook and futur work will conclude this paper.

2 MULTI-MODAL BIOMETRIC SYSTEM

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 serve 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.

2.1 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 (Handmann et al., 2012; Hanheide et al., 2008). For our system a robust, full automatic and real-life face-recognition-based person recognizer is employed (Rabie et al., 2008). The basic technique applied here are Active Appearance models (AAMs) First introduced by Cootes et. al. (Cootes et al., 2001). The generative AAM approach uses statistical models of shape and texture to describe and synthesize face images. An AAM, that 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 (Rabie et al., 2008) 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 (Castrillón et al., 2007). 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 (Hanheide et al., 2008). 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 (Schölkopf and Smola, 2002) is applied as classifier.

2.2 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 online capturing of palm vein images an M2Sys scanner is used. This device uses a near infrared light to create a vein-map of the user’s 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 (Zhang and Hu, 2010) 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, the an edge detecting method is uded 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 (Tong et al., 2012).

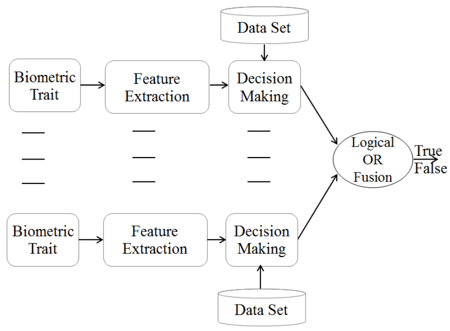


Figure 2: Logical OR fusion method used to combine multiple biometric Traits.

2.3 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 (Miura et al., 2004). 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.

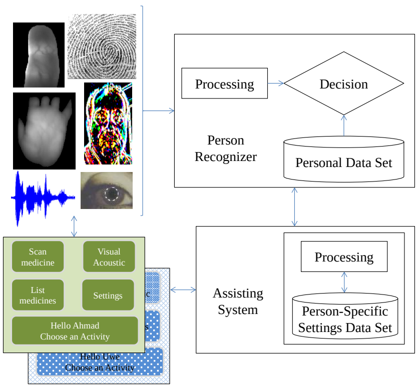


Figure 3: Basic architecture of the medical assisting system. Person is identified by person recognizer, while the assisting system offers person-specific services.

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. I order to detect the direction of the dark line the depth of the valley is scanned with varying angle θi. 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 θi, 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 position 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 means 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 (Miura et al., 2004).

2.4 OR Logic Fusion Method

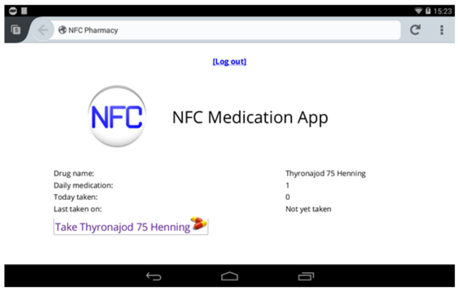
As we strive for just getting the persons living in the home environment identified and our application does not demand very high recognition rate comparing to forensic and boarding applications (Miroslav et al., 2012), it suffices when the person is identified by means of analysis at least one biometric treat. That means, logically, that the decisions of applied stand-alone biometric systems (uni-modals) have to be joined using simple OR rule. To achieve that the face-based, finger-vein-based and hand-palm-vein-based uni-modal sub-systems are joined in one multi-modal person recognition system using a simple OR logic, as depicted in Fig. 2.

3 PERSON-SPECIFIC SERVICE PROVIDER

The basic structure of the whole system could be divided in two basic subsystems, person recognition subsystem and medicine organizer, as illustrated in Fig. 3. The former is based on the analysis of one or more biometric treats in order to identify the interaction partner, while the latter is a PC-tablet or smart phone equipped with modern utilities such as NFC and blue tooth. The system has the role of adapting according to the needs of its interaction partner. After the person is identified by the multi-modal biometric sub-system the assisting system modifies first the system-user interface accordingly. This modification could include the changing of system screening theme and interaction medium. Second the system serves the adapting of the surrounding environment towards the needs of its interaction partner. This could include the setting of heating and light conditions or even giving commands to prepare preferred meals or coffee. Socially the system could call the person-specific contact list and announce the daily arrangements, appointments and activities. Additionally the system sustains the privacy issues of multiple persons living in one household as it has the ability of connecting several portable devices according to the identified ID. Face, finger vein or hand palm vein data are acquired from the elderly 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 glucose meter or by a hand palm vein sensor hidden in a hand air dryer (inh,). 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 serice provider. After the interaction partner is identified the serice provider (PC-tablet, smart phone) 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 settings for the surrounding environment are then called and processed.

4 CASE STUDY: NFC-BASED MEDICAL ORGANIZER

An NFC-Based medical assisting system is implemented as a special case of such assisting systems. The system is implemented on a smart devise, which might be a tablet, smart phone or micro-controller with proper peripheries. The system has the ability of reading and writing NFC tags encapsulated in the medicine packages. Additionally the system is so implemented, that it fetches entries from a pre-saved medicine data base according to the ID of the scanned medicine. The organizer gives then either feedback to the computer system in order to display the doses of the detected medicine or it displays it in its own screen. The organizer gives alarm if the time of the next taking is elapsed or if the next taking of this medicine is taking place prematurely. In practice, the biometric subsystem is implemented on a normal PC equipped with suitable biometric devices, as discussed in chapter 2. A nexus PC-tablet is used for implementing the NFC-based medical organizer. The communication between both subsystems 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 medicine data base, which is saved on the PC-tablet, synchronized permanently. When the interaction partner is identified by the biometric system the ID of this person will be sent to the tablet in order to fetch the corresponding entries of the medicine data set. Assigning MAC addresses of portable devices ensured secure connec-



(a)

(b)

(c)

Figure 4: snapshot of the medical organizer implemented on a PC-tablet. The figure displays the medication of the scanned medicine (a), alert that the the medicine has been already taken (b) and that the taking time of it is due (c).

tion between the biometric system and the only one device of the identified person. That sustains privacy issues of persons sharing one household. The system is evaluated from two subgroups with four persons each. The former was a group of students in the age between 20-25 year, while the latter was a group of elderly people elder than 60 year. The reason of including the youth student is to take the familiarization with modern devices into account. Although the young group has no bodily limitations that prevents them assisting themselves medicinally, it holds such systems as useful for the modern life with much daily activities. The group of elderly people holds it as compensating for weak perception skills. Nevertheless, smooth adapting of the person-specific interface is demanded in order to overcome the ambiguity the of given feedback to the interaction partner.

5 CONCLUSION AND OUTLOOK

In this paper we presented our approach of integrating a person identification and serive provider subsystems in a person-specific medical assisting system for home environment application. As we strive to provide for persons a touch-free assistance, we used some biometric traits of face, finger vein and hand palm vein to identify persons as a prior step towards offering person-specific services. As a special case we presented an person-specific NFC-based medical assisting system, which provide healthy support for elderly and disabled people and offers for them a reliable alleviation for their ordinary life situations. The results indicate that the system was useful for the target group. It provides for them a complete healthy support and relieve their warrens about the accurateness of the taken medicine and the taking time.

An open issue concerns the problem of ambiguity of the given feedback to the interaction partner, as the style of the feedback should depend on the bodily limitations of the interaction partner. To solve this problem we focus for the next step on the considering of the whole health status and the bodily limitations of the interaction partner. Another aspect is the inserting of closed loop health services within the loop, which allows for medical practitioner and pharmacist accessing the medical profile of considered people for control and support reasons (Dohr et al., 2010). Taking the analysis of affective states of the interaction partner as a feed back into account should add a reasonable improvement to the whole performance of the system (Rabie et al., 2009; Rabie and Handmann, 2011). A future comprehensive evaluation with a larger set of test persons could validate the applicability of the system in real life conditions.

ACKNOWLEDGEMENTS

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

REFERENCES

Fraunhofer inhaus zentrum www.inhaus.fraunhofer.de/de/downloads/broschuere.html.

Castrillón, M., Déniz, O., Guerra, C., and Hernández, M. (2007). Encara2: Real-time detection of multiple faces at different resolutions in video streams. *Journal of Visual Communication and Image Representation*, 18(2):130–140.

Cootes, T. F., Edwards, G. J., and Taylor, C. J. (2001). Active appearance models. *Pattern Analysis and Machine Intelligence PAMI*, 23(6):681–685.

Dohr, A., Modre-Osprian, R., Drobics, M., Hayn, D., and Schreier, G. (2010). The internet of things for ambient assisted living. In *Seventh International Conference on Information Technology*.

Handmann, U., Hommel, S., Brauckmann, M., and Dose, M. (2012). *Towards Service Robots for Everyday Environments*, chapter Face Detection and Person Identification on Mobile Platforms, pages 227–234. Springer.

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

Miroslav, B., Petra, K., and Tomislav, F. (2012). Building a framework for the development of biometric forensics. In *35'th International Convention on Information and Communication Technology, Electronics and Microelectronics*.

Miura, N., Nagasaka, A., and Miyatake, T. (2004). Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine Vision and Applications archive*, 15:194–203.

Ouchi, K., Suzuki, T., and Doi, M. (2004). Lifeminder: a wearable healthcare support system with timely instruction based on the user's context. In *8'th IEEE Inter. Workshop on Advanced Motion Control*.

Rabie, A. and Handmann, U. (2011). Fusion of audio- and visual cues for real-life emotional human robot interaction. In *Annual Conference of the German Association for Pattern Recognition*.

Rabie, A. and Handmann, U. (2014a). Nfc-based person-specific medical assistant for elderly and disabled people. In *International Conference on Next Generation Computing and Communication Technologies*.

Rabie, A. and Handmann, U. (2014b). Person-specific contactless interaction, usability of assistance and information systems at home (usahome). *Catalog Human-Machine Interaction, NRW Innovation Alliance, 2014*.

Rabie, A., Lang, C., Hanheide, M., Castrillon-Santana, M., and Sagerer, G. (2008). Automatic initialization for facial analysis in interactive robotics. In *The 6'th International Conference on Computer Vision Systems*.

Rabie, A., Vogt, T., Hanheide, M., and Wrede, B. (2009). Evaluation and discussion of multi-modal emotion recognition. In *The 2'nd International Conference on Computer and Electrical Engineering ICCEE*.

R.A.Ramlee, M.A.Othman, Leong, M., M.M.Ismail, and S.S.S.Ranjit (2013). Smart home system using android application. In *International Conference of Information and Communication Technology ICoICT*.

Ross, A. and Jain, A. (2003). Information fusion in biometrics. *Pattern Recognition Letters*, 24(13):2115–2125.

Schölkopf, B. and Smola, A. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge Press.

Tong, X.-F., Li, P.-F., and Han, Y.-X. (2012). An algorithm for fingerprint basic minutiea feature matching. In *In-*