Face Detection and Person Identification on Mobile Platforms

Uwe Handmann, Sebastian Hommel, Michael Brauckmann, and Michael Dose

Abstract. The Desire project aimed at the development and implementation of a mobile service robotic research platform (technology platform) able to handle real world scenarios regarding service robotic tasks. Different modules for different tasks plus an interaction infrastructure were integrated on this platform. An example of a real world scenario task is the support of a handicapped person to clean up a kitchen in home environments.

One of the main challenges to be solved in this field is the interaction with people. To start an interaction process between a robot and a person, the most important information is the knowledge about the interacting partner's identity and whether the interacting partner is present or not. This means, the robot must be able to detect and be finally able to identify persons. Accurate identification of specific individuals has to be done by analyzing the individual features of each person. A typical feature set that allows for a distinct identification of a specific person is often extracted from the facial image acquired by a camera. This feature-set is stored in a database to allow the identification of different persons independent from place and time by comparing given feature-sets. Thus, a face recognition module was integrated into the technology platform which includes face detection and identification algorithms.

1 Introduction

The technology platform of Desire is equipped with a drive system. Two high resolution cameras are mounted on top of a pan tilt unit (fig. 1). The basic task presented in this paper is the integration and optimization of a person recognition

Uwe Handmann · Sebastian Hommel
Hochschule Ruhr West | Campus, Tannenstraße 43 46240 Bottrop, Germany
e-mail: `uwe.handmann@hs-ruhrwest.de`

Michael Brauckmann · Michael Dose
L-1 Identity Solutions AG, Universitaetsstr.160 44801 Bochum, Germany
e-mail: `MDose@l1id.com`

E. Prassler et al. (Eds.): Towards Service Robots for Everyday Environ., STAR 76, pp. 227–234.
springerlink.com

module on the technology platform using images acquired by the given cameras. This module splits in two parts: a fast face detection component and an accurate face identification component to establish an interaction with known persons.

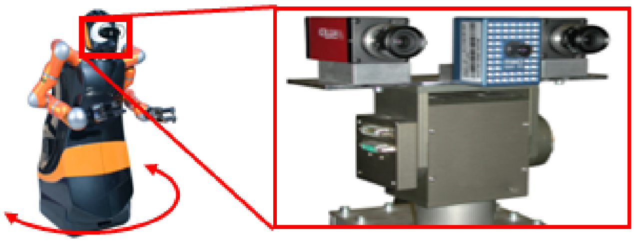

Fig. 1 Mobile technology platform with pan tilt unit and two mounted cameras (red).

2 Face Detection Module

In the service robotic field it is necessary that the robot is able to interact with people in its surrounding area. As a prerequisite the robot has to know in which direction the platform must be moved in order to start the interaction process with someone. In general people could be detected by various kinds of sensors [1][8][9][11]. In the Desire project human detection is performed by a hierarchical analysis using different sensors.

In a first processing step a two-dimensional laser range finder is used to generate hypotheses of human legs in order to position the cameras coarsely for the face detection analysis [2]. In a second processing step a face finding process fine tunes the position of the cameras in order to extract faces from the given camera images with maximum resolution and to track faces found for the repositioning process of the camera over time (fig. 2).

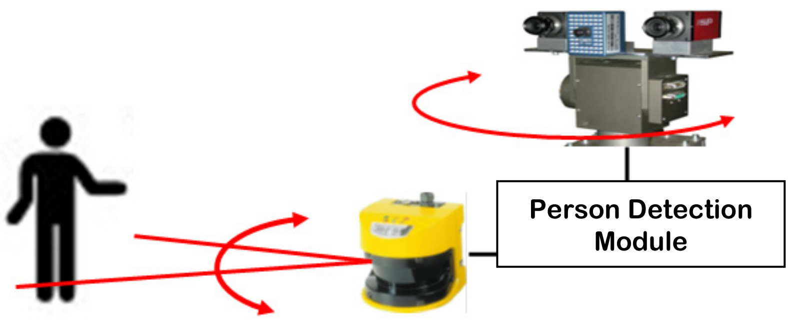

Fig. 2 Laser range finder based leg-detection and vision based face detection combined to a person detection module.

This paper describes the second step of the person finding process. The objective of this part at the beginning of the project was the enhancement of the available face detection algorithms [8][9].

The enhancement of face detection is a two-dimensional problem. First it is necessary to maximize the processing speed of the algorithm in order to find faces at all (fig. 3). The typical person is not cooperative enough to stand still in front of the camera of the interacting robot. Therefore, the face detection algorithm needs to have high speed in order to find and track faces in the given video stream. Also, the repositioning of the camera has to be done before the track moves outside the area observed by the camera, especially if the detected person is walking around.

Secondly, the false alarm rate of the detector has to be very low, i.e. only a very small number of detections that do not contain faces should occur (false positives). In cases of false positives the repositioning of the camera will fail.

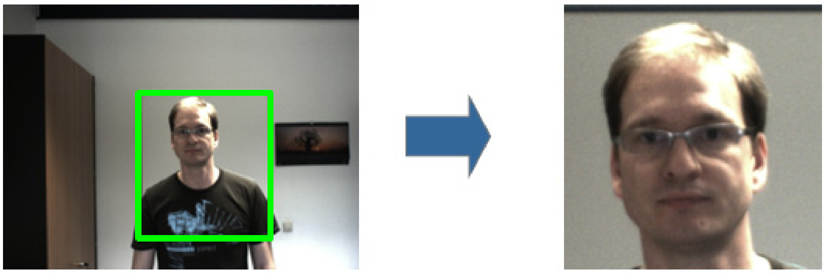

Fig. 3 Image based face finding and face extraction.

In the Desire project a face detection algorithm developed in a preceding project [8] was employed. This algorithm is very accurate in comparison to standard face detection algorithms (e.g. [6][7][5]) and allows fast face detection (frame rate 25 frames/sec) on low resolution images acquired by an analog camera device (image size: 384x288 pixels). The frame rate of this algorithm decreases significantly on high resolution images (image size: 1388x1038 pixels) acquired by a digital video camera.

An acceleration of the frame rate up to 10-12 frames per second was achieved by means of a pyramidal based technique (fig. 4) and a specific color conversion algorithm. With this enhancement it was possible to reposition the camera without losing track of an interacting person.

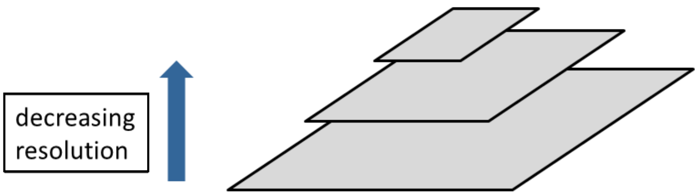

Fig. 4 Pyramidal feature analysis using a specific color space.

3 Face Identification Module

The recognition process typically compares individual feature-sets that are extracted from an acquired image against feature-sets of known people that are stored in a given database. For example, feature-sets, also known as biometric templates, could be fingerprint features such as minutiae or facial features such as texture descriptors. In order to decide if a seen individual is contained in the database, a one-to-many comparison has to be performed. Since there is no claimed identity available for the person in front of the system, the given template has to be compared against all known templates stored in the database, the so called gallery [3][4][10]. The decision whether a person is known to the system or not is made on the most similar gallery element found, and which needs to exceed a threshold. If no template of a given person is stored in the gallery, identification fails, because no template in the database is similar to the acquired template (fig. 5).

In the service robotic environment, the face-identification process is used to recognize an interaction partner. This means that a template of every possible interaction partner has to be stored in the gallery.

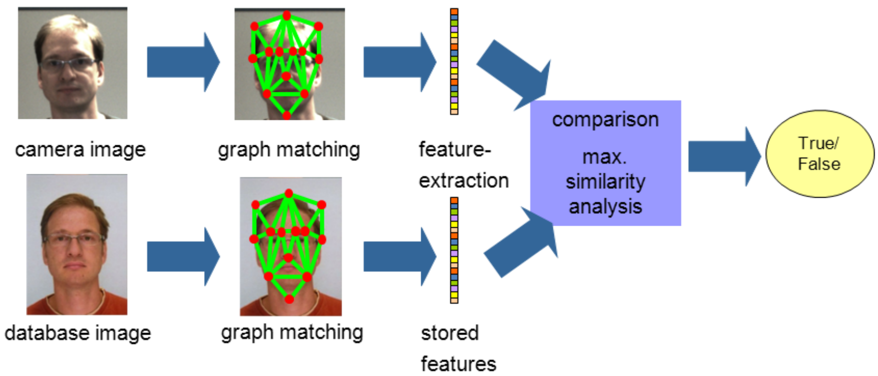

Fig. 5 Face identification process: an acquired image is used to extract a new template. This feature set is compared against templates that are stored in a gallery. Finally a decision using similarities between templates is made.

In the Desire project a face-identification technology based on Hierarchical Graph Matching (HGM) was used [3]. As the name suggests, a critical ingredient of this method is the use of a graph, i.e. a set of nodes and connecting edges, which represents the facial geometry (fig. 6). During the matching process the graph locates itself onto the face so that its nodes converge to predefined salient points (so called landmarks).

The flexibility of the graph further allows an adjustment to size, pose and facial expression. This process runs fully automated thus also solving the face finding problem. Instead of using the raw image pixels, Graph Matching transforms the image information to features more suited for automatic recognition than the grey level information. Graph Matching makes use of specialized kernel functions, which thrive against changes in illumination and contrast. For feature generation a set of these filter functions is applied to each node of the graph, providing a vector of several thousand characteristic descriptors for the face. Several comparison functions may be used to calculate the similarity between two faces. For the analysis of a face, the shape ("landmarks") and the texture ("features") of the face are separated, making HGM a very robust facial recognition method. This technology implies the necessity to detect the best frontal face image of a person over time to be able to extract all landmarks in a similar manner.

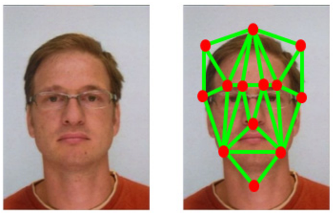

Fig. 6 This figure shows a facial image and the graph placement. The matching process runs fully automated and the graph covers exactly the face. This indicates, that the matching process was successful. Comparing the two images shows that the face has been centered and a correcting in-plane rotation has been applied.

In a non-cooperative scenario it is a difficult task to find images with faces in a frontal view. Therefore the main focus in the Desire project was to enhance robustness of the HGM-technology for non-cooperative face images. In such environments the interacting person usually does not look in the direction of the camera, so in most cases it is nearly impossible to capture face-images in a frontal view. Especially if the interacting person is walking around, often only non-frontal face-images or images with non-optimal illumination could be acquired. To be able to handle these low-quality face images a multi-level architecture was implemented.

On the first level a new skin-texture analyzing algorithm was integrated. The analysis of local skin textures was possible because of the better resolution of the high resolution camera on the technology platform. Using this technique, it was possible to establish multiple local skin analyzers on different regions inside the face (fig. 7). This technique, combined with the existing HGM-algorithm, resulted in a higher stability in the identification process against occlusions and illumination problems.

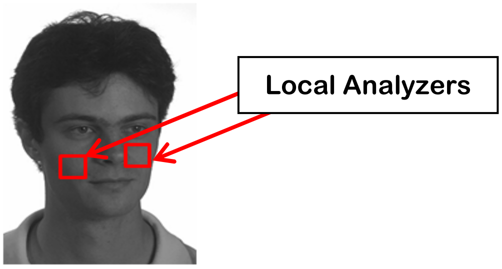

Fig. 7 Local skin analysis enhances face recognition robustness.

On the second level a new pose correction algorithm was integrated. Faces with small poses to the side were corrected in the feature space to a frontal representation (fig. 8).

Fig. 8 Repositioning of landmarks in the feature-space.

On the third level a new algorithm was integrated to be able to handle face-images with large poses to the side. In these face-images the identification process has to deal with occluded areas inside the face (fig. 9).

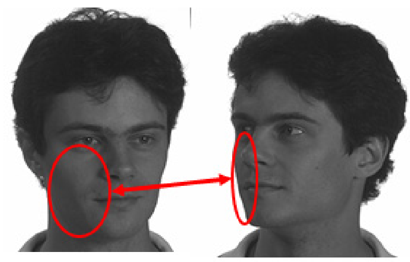

Fig. 9 Occlusion inside a non-frontal face.

On this level all frontal images from the gallery were warped to different non-frontal views using an internal artificial 3D-head representation. Afterwards a feature extraction on the new generated "pose-images" was performed. With this step

Fig. 10 In the middle the acquired image is shown. Left/Right, warped non frontal face images.

comparable representations in the feature-space were generated for the gallery (fig. 10). Finally a decision fusion process was implemented to solve the person identification task.

The resulting face recognition engine was evaluated during internal tests (database with 13 000 images), and adapted versions were examined during external tests, e.g. Face Recognition Vendor Test 2006 [4]. It was shown that the multi-level architecture resulted in significantly higher identification accuracy in the domain of non-cooperative face-images in comparison to the initially used identification technology [4].

4 Summary

In this paper the person recognition module of the Desire project was described. This module was integrated on a mobile service robotic platform to be able to establish an interaction process between a robot and a person. The person recognition module is subdivided in two parts: the person detection part with a camera feedback loop and the face identification part. In both parts an algorithmic enhancement was actualized.

A fast face finding algorithm was integrated in the person detection part and implemented on the technology platform. The feedback loop was used to fuse given information from a two dimensional laser range finder, and results extracted from the face-finding algorithm.

The existing face identification module was upgraded to a multi-layer architecture with respect to non-cooperative scenarios and was also implemented on the technology platform.

Acknowledgments. This work was supported by the German Ministry of Education and Research (BMBF), reference number: 01IMEO1K.

References

[1] Arras, K.O., Martinez-Mozos, O. (eds.): ICRA 2009 Workshop Proceedings of People Detection and Tracking (2009)

[2] Arras, K.O., Lau, B., Grzonka, S., Luber, M., Mozos, O.M., Meyer-Delius, D., Burgard, W.: Range-Based People Detection and Tracking for Socially Enabled Service Robots. In: Prassler, E., et al. (eds.) Towards Service Robots for Everyday Environ. STAR, vol. 76, pp. 235–280. Springer, Heidelberg (2012)

[3] Gehlen, S., Rinne, M., Werner, M.: Hierarchical Graph-Matching, European Patent 01118536.0 (2001)

[4] Phillips, P.J., Scruggs, W.T., O'Toole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., Sharpe, M.: FRVT 2006 and ICE 2006 Large-Scale Results, National Institute of Standards and Technology Gaithersburg, MD 20899, NISTIR 7408 (March 2007)

[5] Rowley, H.A., Baluja, S., Kanade, T.: Neural Network-Based Face Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(1) (January 1998)

[6] Viola, P., Jones, M.J.: Robust Real-Time Face Detection. International Journal of Computer Vision 57(2) (2004)

[7] Viola, P., Jones, M.J.: Fast Multi-view Face Detection, Mitsubishi Electric Research Laboratories, TR2003-096 (August 2003)

[8] Wichert, G.v., Klimowicz, C., Neubauer, W., Wösch, T., Lawitzky, G., Caspari, R., Heger, H.-J., Witschel, P., Handmann, U., Rinne, M.: The Robotic Bar – An Integrated Demonstration of a Robotic Assistant. In: Advances in Human-Robot Interaction. Springer Tracts in Advanced Robotics (STAR), vol. 14. Springer, Heidelberg (2004) ISBN: 3-540-23211-7

[9] Wiegand, S., Igel, C., Handmann, U.: Evolutionary Multi-Objective Optimization of Neural Networks for Face Detection. International Journal of Computational Intelligence and Applications 4(3), 237–253 (2004)

[10] Zhao, W., Chellappa, R., Rosenfeld, A., Philips, P.J.: Face Recognition: A Literature Survey. ACM Computing Surveys 35(4), 399–458 (2003)

[11] Yang, M.-H., Kriegman, D.J., Ahuja, N.: Detecting Faces in Images: A Survey. IEEE Transactions On Pattern Analysis And Machine Intelligence 24(1) (January 2002)