2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

Person Recognition on a Segway Robot: A Video of UT Austin Villa Robocup@Home 2007 Finals Demonstration

W. Bradley Knox, Juhyun Lee, and Peter Stone
Dept. of Computer Sciences
The University of Texas at Austin
{bradknox, impjdi, pstone}@cs.utexas.edu

I. Introduction

This video shows a Segway robot from the University of Texas at Austin competing in the finals of the 2007 Robocup @Home competition, which featured home assistant robots performing various challenging tasks. This demonstration combines a few tasks which will likely be performed by a future home assistant robot. The robot learns a human's appearance, follows the human with his back turned, distinguishes the human from a similarly clothed stranger, and adapts when it notices that the human has changed his clothing. For this task, we introduce a novel two-classifier architecture, using the subject's face as a primary identifying characteristic and his shirt as a secondary characteristic.

A. Robocup @Home

RoboCup@Home is an international research initiative that aims "to foster the development of useful robotic applications that can assist humans in everyday life" [1]. In 2007, its second year of existence, RoboCup@Home attracted eleven custom-built robots from ten different countries and five different continents.

In early rounds, robots performed specific tasks such as following a human around the room, searching for and locating previously seen objects, and differentiating previously seen and unseen humans. Each specific event plus a freeform demonstration was scored and five teams advanced to the Finals. In the Finals, the five finalists performed demonstrations for trustees of the Robocup organization, who determined the final standings. This video shows the demonstration of our team's robot, UT Austin Villa.

B. The Segway Robot

The robot consists of a Segway Robotic Mobility Platform (RMP) 100[1], supporting an onboard computer and various sensors. The Segway provides controlled power in a relatively small package. These are especially helpful in a domestic environment, for it is small enough to maneuver a living environment built for humans and powerful enough to reliably traverse varying indoor terrain including rugs, power cords, tile, and other uneven surfaces.

A tablet PC sits atop the Segway platform, performing all processing onboard. Two cameras and one laser range finder are available to sense the robot's environment. Despite its power, the robot is quite safe, with safety mechanisms such as automatic shut-off, an emergency kill rope, and speed caps at both the hardware and software levels.

II. The Demonstration

Accurate person-recognition will be an essential capability of any fully functional home assistant robot. Asking for identification each time would be cumbersome and unnatural. Instead, the robot should identify the person by context as a human would. This context includes, among other things, visual data, which our algorithm uses.

[1]www.segway.com

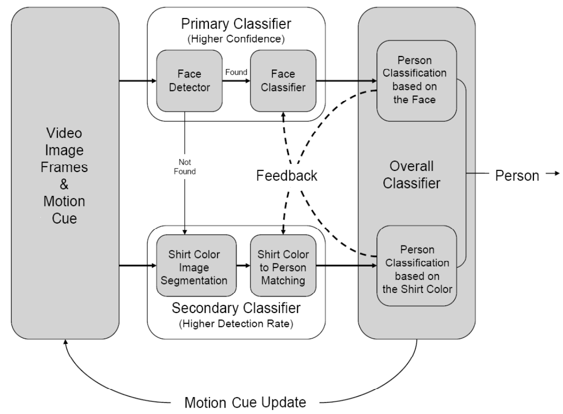

Fig. 1. A conceptual diagram of the person-recognition algorithm used in the Finals demonstration.

Person recognition must be robust. Facial recognition alone is not enough, since humans will sometimes be facing away from the robot's camera. We used shirt color as a secondary classifier to add robustness. In the demonstration, we used shirt detection to identify a person when he turns his back to the robot's camera.

Person recognition, in addition to being robust, must be flexible. Rigidly learning a person's exact appearance at one moment will likely not be sufficient to identify him or her after a significant change in appearance.

Changes in human appearance can be roughly categorized into two types. One occurs quickly, like the changing of clothes every day or cutting one's hair. The other type of change occurs very gradually and includes growing older and losing or gaining weight.

Although we created an algorithm to handle certain cases of both types, the five minute window of our demonstration limited us to creating a scenario that includes only quick changes.

Our scenario was designed to display our algorithm's robustness and adaptability. Specifically, it shows person identification using shirt color as a secondary classifier in the absence of the primary classifier, the face. It also mimics the daily (or so) occurrence of a human changing clothes, showing the robot adapt to this change in the secondary classifier. Lastly, it shows that the Segway robot can effectively follow a recently learned person, given adequate lighting and a shirt color that is distinguishable from the background colors.

III. Primary Classifier

We used the face as the input for our primary classifier. Challenging in its own right, face-recognition becomes even more difficult when performed by a mobile robot. Perspectives change as both robots and humans move. Illumination changes as humans move through different lighting conditions. When the mobile robot is in motion, the camera is, of course, also in motion, downgrading the image quality from its camera(s). Computational limitations are also acute. Faces must be recognized in real time with what computational resources are connected to the robot. A number of successful face-recognition algorithms exist (e.g. [2]), and [3] and [6] describe combinations of face recognition and shirt recognition, but we found none that fit these needs of a home assistant robot.

We detected the person's face in the camera image using Viola and Jones' real-time face-detection algorithm [5]. The given bounding box was used to extract samples from the 50 images. These samples were used as input for our face-recognition algorithm, which learned each person's facial features and later classified unknown faces using the learned features.

Our novel face-recognition algorithm was built on scale-invariant feature transforms (SIFT) features [4]. Given two bounding boxes possibly containing the same face, we extracted SIFT features from both images and counted the number of feature matches (shown in Figure 2). If the matches exceeded a certain threshold, the two objects were considered to be the same object. Our face-recognition method performed well if illumination remained stationary or changed only a moderate amount.

(a) same person (b) different people

Fig. 2. The locations of matched SIFT features are connected with a line.

In other applications of the Segway robot, a classifier was built for each person that the robot learned. Here, there was one known and one unknown person, so only one classifier was used.

IV. SECONDARY CLASSIFIER

A person's shirt was used as a secondary classifier, which effectively supports face recognition because a shirt is more often visible on a human than his or her face. To sample the person's shirt during training, first the algorithm scanned incoming video images for a face, again using Viola and Jones' face detection algorithm. If it detected a face, the bounding box of the face (as given by Viola and Jones' algorithm) was used to choose three other bounding boxes that provided positive and negative shirt samples. One box was drawn directly below the face to provide the positive sample pixels from the shirt. The two other bounding boxes were drawn to the left and to the right of the face. These provided negative sample pixels from the background. An example of the bounding boxes used for shirt training and other examples relevant to the shirt classifier can be seen in 3. Training data consisted of 50 face-containing frames, which took about 10 seconds to collect. Both the positive and negative sample pixels were then inserted into respective histograms. Each histogram was normalized and a final histogram was created by subtracting the negative histogram from the positive histogram. From this final histogram, an RGB color cube was created in which each RGB value contained a boolean value indicating whether or not the RGB value was associated with the shirt color.

Once the training was over, our robot was ready to track and follow the person. The classifier mapped a 320×240 webcam image to a same-sized image with boolean values replacing RGB values

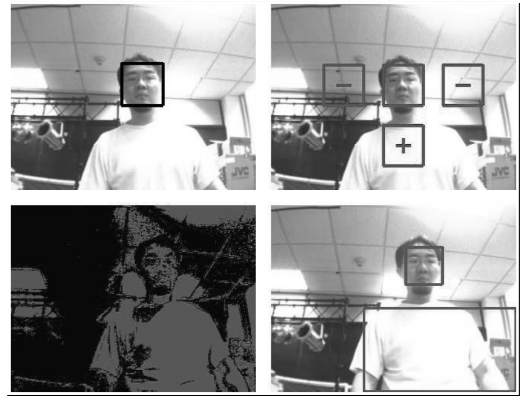

Fig. 3. This sequence of pictures shows the training and testing of the shirt classifying algorithm. Top-left: face-detection; top-right: positive and negative sampling of the shirt; bottom-left: boolean mapping of shirt colors onto image; bottom-right: shirt detection

at each pixel location. After the mapping, the classifier looked for blobs of boolean true pixels in the mapped image. Among many possible shirt candidate blobs, the blob representing the shirt was chosen by its size and its proximity to recent shirt blobs.

V. INTERCLASSIFIER FEEDBACK

In this demonstration the secondary classifier was enabled only when the robot did not detect a face. While the primary classifier was being used and the secondary classifier was disabled, the shirt pixels below the detected face were used to check the accuracy of the shirt classifier. If the two classifiers disagreed (e.g. the face is classified as negative and the shirt is positive), the secondary classifier was deleted and retrained. A conceptual diagram of the algorithm is shown in Figure 1.

VI. CONCLUSION

The demonstration went smoothly. We only noticed one flaw, which was cut from the video. At one point the robot mistook a mahogany set of shelves for the red-shirted human and turned away from the human. As this error displayed, a flaw of our shirt classifier is that its color resolution is much coarser than that of the human vision system, allowing similar colors to look identical to the shirt classifier.

The panel of judges scored the finalists, determining each team's final standing in Robocup@Home 2007. We received second place, behind the AllenamaniACs from RWTH Aachen in Germany.

Future work should continue to seamlessly integrate different capabilities of the robot, aiming to create a assistive agent that interacts naturally and safely in a home environment.

For more information on our Segway robot, please visit our team website at `http://www.cs.utexas.edu/ AustinVilla/?p=athome`.

REFERENCES

[1] Robocup@home rules and regulations. http://www.ai.rug.nl/robocupathome/documents /rulebook.pdf.

[2] P. Belhumeur, J. Hesphana, and D. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *PAMI*, 1997.

[3] M. Everingham, J. Sivic, and A. Zisserman. Hello! my name is... buffy – automatic naming of characters in tv video. *MBVC*, 2006.

[4] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 2004.

[5] P. Viola and M. Jones. Robust real-time face detection. *IJCV*, 2004.

[6] S. Waldherr, R. Romero, and S. Thrun. A gesture based interface for human-robot interaction. *Autonomous Robots*, 9(2):151–173, 2000.