A Robot Capable of Detecting, Identifying, and Tracking a Human Individual

Takashi Koyama[1,#], Takuya Nakata[2,#], Masaaki Iiyama[3*] & Michihiko Minoh[3]

[1]Todaiji Gakuen Senior High School, [2]Osaka Prefectural Kitano Senior High School, [3]Academic Center for Computing and Media Studies, Kyoto University

Abstract

We have developed a robot that can track a human individual using computer vision technology. The robot, which has both a consumer camera and an omnidirectional camera mounted on its body, can detect a person using the cameras and subsequently track him/her. The use of moving cameras poses problems that do not affect static cameras. This paper addresses the following two problems: situations in which more than one person is present in front of the camera, and situations that the robot loses sight of a person. We propose a method to solve these problems, and experimental results with our robot show that our method successfully tracks only the target person, even if other persons are present, or if the person is out of sight.

Key words: Person detection, Facial detection, PID controller, Omnidirectional camera

1. Introduction

The performance of computers and digital cameras continues to improve. This trend motivates active development in the field of computer vision. The principle of computer vision is to use computers equipped with cameras to analyze various phenomena in the real world. Among various technologies of computer vision, one of the most important is human tracking, which keeps track of the locations of humans in real time using images obtained from cameras. Human tracking is a key component for making computers understand human behaviors and to cooperate with humans. In particular, for surveillance, health care, and service industry applications, it is highly beneficial to use computers to observe the behavior of humans, such as criminals, patients, and customers. The aim of this work was to make full use of modern human-tracking technology to develop a robot that can detect and track a person in an indoor environment. Additionally, the robot had to follow and record the behavior of the target. To facilitate constant observation of the targeted person, we equipped the robot with a camera.

By mounting the camera on the robot, the robot can track the targeted person while moving freely. This flexibility, however, can cause two serious problems in real environments. First, when there is more than one person present, it is not obvious to the robot which person it should track. Second, when the targeted person leaves the sight of the robot, it must find and look at his or her face again. Otherwise, it cannot continue tracking the target. One potential solution to this problem is to mount more than one camera on the robot to cover all directions. In practice, however, using multiple cameras to cover all directions is both expensive and computationally demanding. Hence, we attempted to address this problem without increasing the number of cameras.

To overcome these two problems in robot-based human tracking, we proposed two novel approaches. First, we developed an algorithm of human tracking that can distinguish a target from others. Second, we introduced a special camera for robot-based human tracking, thereby achieving robustness even when a target is out of sight.

2. Robots for Robust Human Tracking

To implement human tracking in our robots, we use *Ubuntu*, a free distribution of *GNU/Linux* for the operating system of the robots, and *Python 2* as a programing language. For human detection, we employed algorithms provided by *OpenCV*[(1)], an open-source library for computer vision. For robot hardware, we used Roomba, which is an automatic cleaner for home use. To provide the robots with enough computational power for human tracking, we installed a laptop computer on the top of the Roomba.

To reiterate, we have two objectives. One is to distinguish a target person from others. The other is to infer the direction of an out-of-sight target. To achieve these, we created two robots, one to address each of the two problems.

2.1. A Robot that Can Distinguish a Human Target

For the robot that can distinguish a particular human target from other individuals, we used a consumer camera. To control the robot, we used a proportional-integral-differential (PID) controller, which enables us to achieve accurate tracking with simple programming and parameter tuning.

To begin, we attempted to distinguish individuals using facial features obtained from camera images, which is commonly used in image-based human detection[(2)]. However, this approach requires that the camera always capture a frontal view of the target, and that the captured images be clean, i.e., not affected by degradations such as motion blur. As we found that it is difficult to satisfy these requirements using the camera of the moving robot, we decided not to rely on facial features, but instead to only use the positions of faces, as described below.

Suppose that we are tracking a person with our robot, and have captured an image frame illustrated in the upper part of Fig. 1, where only the target person is present in the robot's field of view. After an interval of some time, we capture another frame shown in the lower part, in which we observe three faces. Two of the three faces correspond to persons other than the target, who entered the field of view after the first frame was captured. In addition, the face drawn by dotted lines indicates the previous position of the target's face.

To distinguish the target person from others in the second frame, we calculated the distance between the position of each face and the

#T. Koyama and T. Nakata have contributed equally as first authors.

*Correspondence Researcher: iiyama@mm.media.kyoto-u.ac.jp

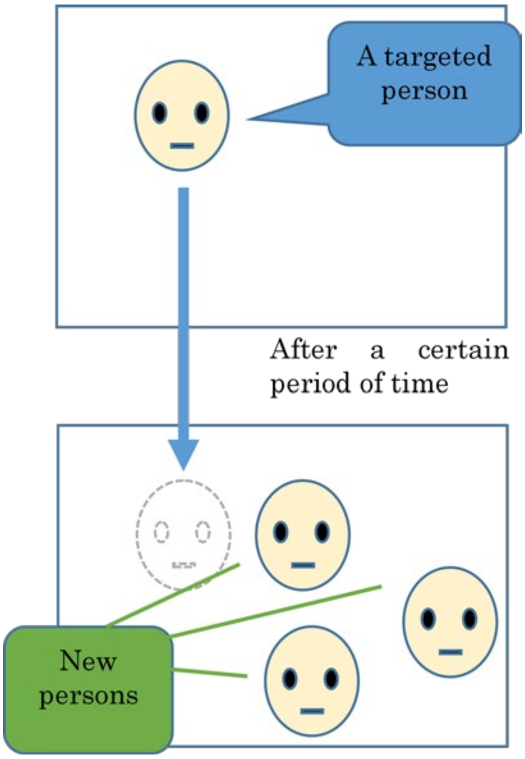

Fig. 1. Finding the corresponding face.

previous position of the target's face, and associated the target with the face that is the nearest to the previous position. In Fig. 2, the face indicated with a red arrow is the nearest to the previous position; thus, our method could determine that this face corresponds to the target.

2.2. A Robot that Can Infer the Direction in Which an Out-of-Sight Human Target is Located

To track an out-of-sight target, we mounted an omnidirectional camera on the robot in addition to the consumer camera. We set the consumer camera in front of the robot and used it for facial detection, and utilized an omnidirectional camera for moving body detection. The consumer camera has a limited field of view, so the system cannot recognize a human when the human is out of sight. Instead of this disadvantage, moving object detection with the omnidirectional camera, which is faster than facial detection, is used. This omnidirectional camera acts like the peripheral vision of human eyes in that neither can distinguish colors but both are good at finding moving objects.

The robot alternately acquires images from either of the two cameras for a few seconds each. If it succeeds in facial detection using the consumer camera, which informs the robot about the coordinates of the target in the image, the robot turns toward the target. However,

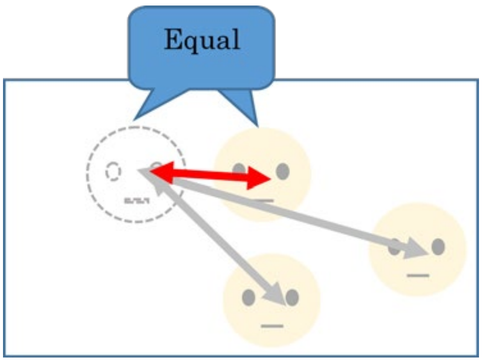

Fig. 2. Nearest face.

if it cannot find any persons using facial detection, the system tries to find moving objects from the image captured by the omnidirectional camera, and subsequently turns toward the detected object.

We used frame difference for moving object detection because cameras move together with the robot. To infer the direction of the out-of-sight person, we used the following method for moving object detection. First, the robot calculates differences around its horizontal direction, and the magnitude of these differences. The magnitude is calculated from the weighted sum of differences in each pixel. Second, the direction with the largest difference is selected. This direction is chosen from 16 predefined directions. Finally, the robot rotates toward this direction.

3. Experiments and Results

Through experiments, we evaluated our two robots for human tracking. All the experiments were conducted in an indoor environment.

3.1. The Robot that Can Distinguish a Human Target From Other Individuals

In the first experiment, we evaluated the robot that can distinguish a targeted person as follows. Fig. 3 shows the robot tracking the target. First, one subject designated as the target stood in front of the robot and allowed the robot to remember the position of his face. Next, another subject approached the designated target. At this point, there were two persons in the robot's field of view. We subsequently observed whether the robot could continue tracking and distinguishing the target from the other person, and additionally, keep him in the center of the field of view. Any movements on the part of the robot resulted in tracking failure, and thus the robot wrongly regarding the second person as the target. Hence, the robot did not move at all when the second person entered the field of view. We therefore concluded that the robot was able to successfully track the target and distinguish him perfectly from the other person.

Our current algorithm does not consider the camera's ego-motion. We supposed that the robot would fail to track the person when the robot itself moves or rotates. Fixing such a situation by estimating and adjusting the robot's ego-motion is a future aim.

3.2. The Robot that Can Infer the Direction of a Person

To evaluate how accurately the robot is able to track out-of-sight targets, we performed two experiments. First, we experimented in a space where there were few moving objects except for the target. The robot almost succeeded in tracking the target even though the target frequently avoided standing in front of the robot. If the out-of-sight target stopped completely, however, it was difficult to track the target. Second, we put the robot in a place surrounded by some moving objects. At first, the robot responded to the objects. The robot was able to eventually track the correct target, but it spent less time if the robot ignored the objects. Fig. 4 shows the robot that lost the person and found him using the omnidirectional camera. In order to improve its tracking performance, it is necessary for the robot to detect faces more quickly. From the above results, we conclude that the robot can adequately track the out-of-sight target using only an omnidirectional camera for moving body detection.

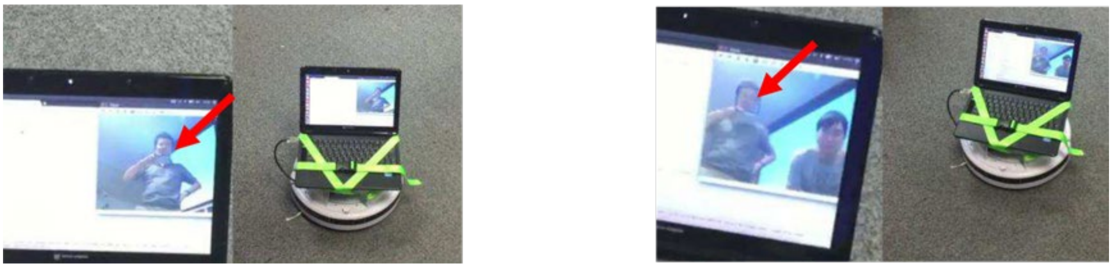

Fig. 3. Robot and captured image. Left: a situation with a single person in the robot's field of view. Right: a situation with two persons in the same field. Detected faces are indicated with red arrows.

Fig. 4. Robot and captured image. Left: the robot detects the person in front of it. Middle: the robot fails to detect the person using its main camera and instead detects moving object with its omnidirectional camera. Right: the robot rotates toward the person.

4. Conclusion

In this work, we proposed novel implementations of human-tracking robots to overcome two problems that arise in real-world environments by adopting different approaches for each problem. The first robot can distinguish a target person from others robustly using facial positions, and continue tracking the target even if more than one person is present in its field of view. On the other hand, the second robot can infer the direction where the target is located even if it loses sight of the target.

These robots have room for improvement. Our robots cannot continue tracking when a target turns his or her face away from them. While our current implementations rely only on facial detection for tracking, we will need to introduce other types of information to our human tracking algorithms to deal with situations where we cannot obtain the frontal views of faces. Another issue is that our robots cannot follow a person when he or she makes sudden movements. Finally, we have yet to integrate our two approaches proposed here into a single robotic framework. In combination, they will enable more robust robot-based human tracking that can handle a wide range of challenging scenes in the real world.

Acknowledgments

We would like to express gratitude to tutors of the Minoh laboratory for supporting our experiments. We also thank the ELCAS program for providing this valuable opportunity.

References

1. OpenCV. <http://opencv.org> [accessed 16 Jul. 2016]. (2000 onwards).
2. Viola, P. & M. Jones. Robust Real-time Object Detection. Int. J. Computer Vision 57: 137–154. (2004).

人物を捜索・特定・追跡するロボット

小山 高[1,#]，中田 匠哉[2,#]，飯山 将晃[3*]，美濃 導彦[3]

[1]東大寺学園高等学校，[2]大阪府立北野高等学校，[3]京都大学学術情報メディアセンター

要旨

我々は，画像認識の技術を使った人物を追跡するロボットを製作した．このロボットはカメラを車体の上に取り付けており，カメラで人物を検出し，その人物を追いかける．移動できるカメラを使うことによって，従来の固定カメラでは起きなかった問題が発生した．本論文では，複数の人物がカメラの視野に存在している時や，ロボットが人物を見失った時の解決方法を中心に扱う．これらを解決することで，複数の人物がいても常に対象となる人物のみを追跡し，あらゆる方向から人物の存在する方向を推定するロボットを製作することに成功した．

重要語句：人物検出，顔検出，PID 制御，全周囲カメラ

小山高と中田匠哉は貢献度の等しい筆頭著者である．
* 内容に関する連絡先：iiyama@mm.media.kyoto-u.ac.jp