We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
**Abstract**

In this chapter, we proposed three assistive systems for visually impaired individuals based on image processing: Kinect cane system, Kinect goggle system, and light checking system. The Kinect cane system can detect obstacles of various sizes and also recognize objects such as seats. A visually impaired user is notified of the results of detection and recognition by means of vibration feedback. The Kinect goggle system is another type of wearable system, and can make user’s hands free. The light checking system is implemented as an application for a smartphone, and can tell a visually impaired user the ON/OFF states of room lights and elevator button lights. The experimental results demonstrate that the proposed systems are effective in helping visually impaired individuals in everyday environments.

**Keywords:** assistive system, Kinect, cane, goggle, smartphone, camera, image processing, obstacle detection, object recognition, light checking

**1. Introduction**

The world health organization estimated the number of visually impaired individuals to be approximately 285 million in 2014 [1]. Many of them use white canes to detect obstacles around them. However, the detectable ranges of white canes are very short. Guide dogs are also used to navigate visually impaired individuals to their destinations while avoiding obstacles. However, it is difficult to provide the sufficient number of guide dogs due to long-time periods and expenses to train them. In order to overcome these problems, extensive research has been dedicated to creating assistive systems for the visually impaired [2].

Obstacle detection is one of the representative research themes. Many research groups have proposed obstacle detection systems based on laser sensors [3–11], single charge-coupled
devices cameras [12–16], ultrasonic sensors [17–29], stereoscopic cameras [30–41], or RGB-D cameras [42–46]. These assistive systems are built on the basis of the concept of the electronic travel aid (ETA) [47, 48], which aims to assist visually impaired users in walking while avoiding obstacles. Therefore, these systems can notify the users about obstacles but cannot tell them what kind of objects they are.

Here, let us consider a situation where there is a seat (bench) in front of a visually impaired user as shown in Figure 1, and the user wants to sit on the seat to take a rest. In this situation, the seat is not just any obstacle, but a useful equipment. If the user uses one of the obstacle detection systems mentioned earlier (see Figure 1(a)), he or she is required to confirm the obstacle by himself or herself. However, if the user uses an object recognition system, which can determine the object to be a seat (Figure 1(b)), the user can obtain a benefit. It is necessary to build an assistive system to recognize objects around a visually impaired user.

Several research groups have proposed object recognition systems. Drug packages [49], classroom doors [50–53], podiums [50], and pathways [54, 55] are recognized by using barcodes [49, 56], augmented reality markers [50, 52–54], radio frequency identification tags [23, 51, 57, 58], Bluetooth devices [59], wireless network devices [55], or visible light communication devices [60, 61]. These physical devices are useful, but it is difficult to deploy such devices in everyday environment.

Other research groups have also proposed assistive systems to notify visually impaired users about tables [62], color blocks [63], and staircases [64–67] by means of laser range sensors [65–67] and Kinect sensors [62, 63]. These systems are useful, but are not sufficient yet. Other types of object should be recognized to help visually impaired individuals more.

This chapter proposes our assistive systems not only to detect obstacles of various sizes but also to recognize objects of various types by use of image processing technique.

**Figure 1.** Obstacle detection system (a) versus object recognition system for the visually impaired (b).
2. Kinect cane system

Figure 2 shows our Kinect cane system composed of a white cane and a backpack [68]. A Kinect sensor, a numeric keypad, and a tactile device are attached to the white cane. Kinect is an infrared-based range sensor for a consumer game machine, that is, Microsoft Xbox, and the white cane is also a commercial product for the visually impaired. (The sensor and the cane are approximately 300 and 100 USD, respectively.) The Kinect sensor is set at 75 cm from the floor. These devices are connected by wires with a portable personal computer and a UPS battery in the backpack. The computer and the UPS battery are used for device controlling and power supply, respectively.

In this system, the X, Y, and Z axes of the world coordinate system are defined to be the horizontal vector oriented from left to right, the vertical vector oriented from top to bottom, and the horizontal vector extending from a Kinect sensor into the environment, respectively.

The Kinect cane system can detect obstacles and recognize several objects, such as seats, by use of our special computer programs implemented for the following methods.

2.1. Obstacle detection

The Kinect cane system can detect obstacles which would prevent a visually impaired user to walk safely [45]. We provide two detection methods of small and large obstacles considering
the property of a Kinect sensor. These two methods are simultaneously executed in the obstacle detection mode of the Kinect cane system, and if one of these methods detects obstacles, the tactile device returns vibration feedback to the user.

2.1.1. Detection of small obstacles

Figure 3(a) shows an image of a corridor scene including a small obstacle (i.e., a box with a height of 7 cm) on a floor. Figure 3(b) shows the depth data of the corridor scene. The scene
image and depth data were obtained by the Kinect sensor. In the depth data, the distances from the Kinect sensor to points on object surfaces are coded. Black pixels represent positions where the sensor cannot measure the distances. The colored pixels (except black ones) are called edges in this chapter.

In order to detect a small obstacle on a floor, the method uses a measurement line, which is a vertical line at the center of depth data as shown in Figure 3(b). The measurement line is projected onto the floor plane, and the projected line is called a floor line. Figure 3(c) shows the profile of the depth data along the floor line. Black dots represent edges on the Z-Y plane.

The floor line is formulated by

\[ y = az + b, \]

where \( a \) and \( b \) are coefficients. Let \((z_i, y_i)\) denote the coordinate of the \( i \)-th edge (\( i = \{1, 2, \ldots, I\} \)). The line is fitted to the edges by minimizing the following sum of the squared distances:

\[ S = \sum_{i=1}^{I} (d_i)^2, \]

where

\[ d_i = \frac{az_i + b - y_i}{\sqrt{1 + \left(\frac{dz_i}{dy_i}\right)^2}}. \]

The optimal values of \( a \) and \( b \) are obtained by using a robust estimation [69] as follows:

\[
\begin{pmatrix}
  a^* \\
  b^*
\end{pmatrix}
= \left( \frac{\sum w_i z_i^2 \sum w_i z_i}{\sum w_i z_i} \right) \left( \frac{\sum w_i z_i y_i}{\sum w_i y_i} \right),
\]

where weight values are defined as

\[ w_i = \frac{d_i}{1 + \left(\frac{dz_i}{dy_i}\right)^2}. \]

The equation

\[ y = a^* z + b^* \]

represents the optimal floor line. The edges above the optimal floor line are determined to be the obstacle.

2.1.2. Detection of large obstacles

The detectable range of a Kinect sensor is from approximately 40 to 600 cm. In many cases, the upper limitation, 600 cm, is sufficient for obstacle detection. However, the lower limitation, 40 cm, may cause problems that visually impaired individuals collide with obstacles in front of
them. In this section, we propose a detection method of large obstacles not only in the detectable ranges but also nearer than the lower limitation.

In the method, three small circular windows, called obstacle measurement (OM) spots, are set on depth data as shown in Figure 4. They are arranged horizontally with a certain interval. The positions and interval are determined considering the height and width of the body of a user. Another small circular window, called a floor measurement (FM) spot, is set on the bottom area. The OM and FM spots are represented as $S_{O1}$, $S_{O2}$, $S_{O3}$, and $S_F$, respectively.

If each measurement spot includes enough number of edges, the spot is defined to be detected, and the mean depth value is calculated from the depth values of the edges in the detected spot.

The system determines the distance between a Kinect sensor and a large obstacle on the basis of their relation as follows (see Figure 5):

Case 1: If the obstacle is nearer than 40 cm, all the spots would not be detected. In this case, the system determines the distance to be less than 40 cm.

Case 2: If the obstacle is between 40 and 600 cm, at least one of the OM spots would be detected, and the FM spot would be detected as well. The system outputs the minimum value among the mean depth values of the detected OM spots.

---

**Figure 4.** A large obstacle (pillar) in a building. (a) Color image; (b) depth data.
Case 3: If the obstacle is further than 600 cm, all the OM spots would not be detected, whereas the FM spot would be detected. The system determines the distance to be more than 600 cm.

The pillar was successfully detected in Figure 4.

### 2.2. Object recognition

The Kinect cane system can recognize several objects from depth data. The recognition methods and results are described below.

#### 2.2.1. Planes

Artificial environments are generally composed of many planes, such as floors and walls, and therefore, planes can be effective clues to recognize the environments. Figure 6(a) and (b) shows an example scene and its depth data, respectively. Planes are recognized [70] by using the following method based on random sample consensus (RANSAC) algorithm [71]:

1. Three edges are randomly chosen from edges in depth data, and then a plane is fit to the chosen edges by use of the least-square method. The three pink points in Figure 6(c) and (d) are the randomly-chosen edges, and the blue regions in Figure 6(e) and (f) are the planes.

2. The method determines edges whose distances to the plane are nearer than a threshold. These edges are called *inliers*.

3. Steps (1) and (2) are iterated a certain number of times.

4. The plane with the most inliers is selected. Figure 6(g) and (h) shows the selected plane, which corresponds to the floor. The inlier edges are eliminated.

<table>
<thead>
<tr>
<th>Relation between a Kinect sensor and a large obstacle</th>
<th>Case 1: Less than 40cm</th>
<th>Case 2: Between 40 and 600 cm</th>
<th>Case 3: More than 600cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstacle measurement spot</td>
<td>Not detect</td>
<td>Detect</td>
<td>Not detect</td>
</tr>
<tr>
<td>Floor measurement spot</td>
<td>Not detect</td>
<td>Detect</td>
<td>Detect</td>
</tr>
</tbody>
</table>

Figure 5. Relation between a Kinect sensor and a large obstacle.
5. Steps from (1) to (4) are iterated until the number of the remaining edges is less than a threshold.

In Figure 7, the floor and the two walls were recognized correctly. Black pixels are the remaining edges, which are used for seat recognition described below.

![Figure 6. Processes of recognition of planes (a)-(h).](image)

![Figure 7. Recognition result of planes.](image)
2.2.2. Seats

The sitting surfaces of seats (such as chairs, stools, and benches) are considered to be the most essential parts of the seats and are recognized as regions that satisfy the following conditions:

1. Candidate regions are composed of the remaining edges which are between 30 and 50 cm from the floor.

2. The areas of the candidate regions are more than $1200 \text{ cm}^2$.

Figure 8 shows the recognition result of the seat in Figure 6(a). Red pixels represent the sitting surface of the seat. Figure 9 shows the color image, depth data, and recognition result of other seats. All the seats were recognized correctly.

Figure 8. Recognition result of a seat.

Figure 9. Recognition result of other seats. (a) Color image; (b) depth data; (c) recognition result.
The seat recognition method was applied to 62 sample scenes including seats, and 88% of the scenes were recognized correctly.

It is difficult for this method to recognize seats that have nonparallel sitting surfaces, but there would be not so many such seats in general environment.

2.2.3. Other objects

The Kinect cane system can also recognize upward staircases, downward staircases, and elevators on the basis of the recognition results of planes. The recognition methods are described in detail, for example, in [68, 70], and, in this section, the recognition results are shown in Figures 10–12.

It is difficult to recognize upward staircases composed of only one or two steps, slopes, and elevators with sealed doors. The Kinect cane system is not designed to detect holes. A user can detect holes by using the system as a conventional white cane.

2.3. User interaction

Ordinarily, a visually impaired user can use the Kinect cane system as a conventional white cane. Figure 13(a) shows an example situation where a user walks in an elevator hall. The user has been here several times, and therefore the user knows there is a bench in the hall, but does not know (or forgot) its accurate location. The user stops walking for safety and then instructs the system to find the bench (seat) as shown in Figure 13(b). The user executes the seat

![Upward staircases](image1)

(a)  (b)

Figure 10. Recognition result of upward staircases. (a) Color image; (b) recognition result.
recognition program by pushing the corresponding key on the numeric keypad. The user makes the system search for the bench (Figure 13(c)). If the sensor finds the bench, the tactile device returns vibration feedback to the user (Figure 13(d)). The user walks toward the bench (Figure 13(e)), and then confirms it (Figure 13(f)). Finally, the user can sit on the bench (Figure 13(g)).
Figure 13. A visually impaired user wants to sit on a bench to take a rest in an elevator hall. (a) A visually impaired user comes out of an elevator. (b) The user instructs the system to find a bench. (c) The user pans the Kinect sensor. (d) The system finds the bench. (e) The user walks toward the bench. (f) The user confirms the bench. (g) The user can sit on the bench.
3. Kinect goggle system

This section introduces another type of Kinect-based wearable assistive system, a Kinect goggle system [72] (Figure 14). A Kinect sensor is attached to a goggle on the face of a visually impaired user. A notebook computer, a numeric keypad, a tactile device, and a UPS battery are set in a shoulder bag. These devices are connected with wires for device controlling and power supply. The Kinect goggle system does not require a visually impaired user to hold a heavy Kinect sensor, and therefore the user can make his or her hands free.

The current system can detect obstacles and recognize seats by use of the software which is almost the same as those of the Kinect cane system. Figure 15 shows an example scene including a bench. The red region in Figure 15(c) represents the sitting surface of the bench.

Figure 14. Our Kinect goggle system.

Figure 15. Recognition result of a seat by the Kinect goggle system. (a) A bench; (b) depth data; (c) recognition result.
4. Smartphone-based light checking system

Currently, many visually impaired individuals become to use smartphones as well as sighted people, and therefore it is useful to build assistive systems based on smartphones. In this section, we propose a notification system of the ON/OFF states of room lights and elevator button lights by use of a smartphone camera [73].

4.1. Room light

It is said that many visually impaired individuals often use room lights for the prevention of crime, and so on. They cannot confirm (or have difficulties in confirming) the ON/OFF states by seeing the lights, but can change the states by using the switches of the lights.

There are two types of room light switch. Ones are switches that change their shapes according to the ON/OFF states. The others are switches that do not change their shapes. Figures 16 and 17 show examples of the former and latter switches, respectively.

Visually impaired individuals can confirm the ON/OFF states if the switches are the former types, but cannot if the latter. It is necessary to help visually impaired individuals know whether room lights are ON or OFF. This section proposes a notification system of the ON/OFF states of room lights based on interaction between a visually impaired user, a smartphone, a room light, and its switch.

First, a visually impaired user turns a smartphone camera toward a room light and then changes the switch of the light as shown in Figure 18. The system determines the luminosity

![Figure 16. Seesaw-type switch. (a) ON; (b) OFF.](image1)

![Figure 17. Push-type switch. (a) ON; (b) OFF.](image2)
change of the light by analyzing the intensity histogram and the exchangeable image file format (EXIF) information of an image obtained by the camera. If the light becomes brighter and then darker, the smartphone system will tell the user that the room light turns ON and OFF, respectively. If the user fails to set the smartphone camera, the system returns nothing. Therefore, the user can be aware of the failure and do again after resetting the camera.

In a lecture room shown in Figure 19, a user used the proposed system and was able to correctly determine that the room lights were turned off.

Figure 18. A visually impaired user tries to confirm the ON/OFF state of a room light.

Figure 19. Recognition result of the state of room lights in a lecture room. (a) ON; (b) OFF.
4.2. Elevator button light

The notification system can be also used to confirm the arrival floor of an elevator cage. The light of a floor button on a control panel in an elevator cage often turns off when the cage arrives at the corresponding floor. By using the system, a visually impaired user can know whether the cage arrives at a desired floor as described below.

First, a visually impaired user sets a smartphone camera toward the button of a desired floor as shown in Figure 20 and executes the light notification program. When the cage arrives at the floor, the light of the button will be turned off, and the system tells the user that the light is turned off. By hearing the message, the user can know the current floor.

In Figure 21, a user can correctly determine that he arrived at the ninth floor.

![Figure 20](image)

**Figure 20.** A visually impaired user tries to confirm the current floor.

![Figure 21](image)

**Figure 21.** Recognition result of the state of an elevator button light. (a) ON; (b) OFF.
5. Conclusion

In this chapter, we proposed three assistive systems for visually impaired individuals based on image processing. The Kinect cane system can detect obstacles of various sizes and also recognize objects such as floors, walls, seats, upward staircases, downward staircases, and elevators. The detection and recognition results are notified to a visually impaired user by means of vibration feedback. The Kinect goggle system is another type of wearable system and can make user’s hands free. The system can also detect obstacles and recognize objects. The smartphone-based light checking system can inform a visually impaired user about the ON/OFF states of room lights and elevator button lights. The user can confirm light states and arrival floors. The experimental results demonstrate that the proposed systems are effective in helping visually impaired individuals in everyday environments.

Acknowledgements

The authors wish to acknowledge Dr. Nobuo Ezaki, Professor of National Institute of Technology, Toba College and Dr. Shinji Mizuno, Professor of Aichi Institute of Technology, for advice on design of this study.

This work was supported in part by the JSPS KAKENHI Grant Number 16K01536.

Author details

Hotaka Takizawa1* and Mayumi Aoyagi2
*Address all correspondence to: takizawa@cs.tsukuba.ac.jp
1 University of Tsukuba, Tsukuba, Japan
2 Aichi University of Education, Kariya, Japan

References


[34] Kawai Y, Tomita F. A supporting system for visually impaired persons to understand three-dimensional visual information using acoustic interface. In: Proceedings of the 16th International Conference on Pattern Recognition. The International Association for Pattern Recognition, Quebec, Canada. Vol. 3; 2002. p. 974-977


[48] Hersh MA, Johnson MA. Assistive Technology for Visually Impaired and Blind People. Springer; 2008


for all 7th international conference on Universal access: theoretical perspectives, practice, and experience, ERCIM’02. Berlin, Heidelberg: Springer-Verlag; 2003. p. 409-421


