sensors

MDPI

Article

A Spot Reminder System for the Visually Impaired Based on a Smartphone Camera

Hotaka Takizawa [1,*], Kazunori Orita [1], Mayumi Aoyagi [2], Nobuo Ezaki [3] and Shinji Mizuno [4]

[1] University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8573, Ibaraki, Japan; kc_orita@pr.cs.tsukuba.ac.jp
[2] Aichi University of Education, 1 Hirosawa, Igaya, Kariya 448-8542, Aichi, Japan; maoyagi@auecc.aichi-edu.ac.jp
[3] National Institute of Technology, Toba College, 1-1 Ikegami, Toba 517-8501, Mie, Japan; ezaki@toba-cmt.ac.jp
[4] Aichi Institute of Technology, 1247 Yachigusa, Yakusa, Toyota 470-0392, Aichi, Japan; s_mizuno@aitech.ac.jp
* Correspondence: takizawa@cs.tsukuba.ac.jp; Tel.: +81-29-853-5503

Academic Editor: Vittorio M. N. Passaro
Received: 27 October 2016; Accepted: 23 January 2017; Published: 4 February 2017

Abstract: The present paper proposes a smartphone-camera-based system to assist visually impaired users in recalling their memories related to important locations, called spots, that they visited. The memories are recorded as voice memos, which can be played back when the users return to the spots. Spot-to-spot correspondence is determined by image matching based on the scale invariant feature transform. The main contribution of the proposed system is to allow visually impaired users to associate arbitrary voice memos with arbitrary spots. The users do not need any special devices or systems except smartphones and do not need to remember the spots where the voice memos were recorded. In addition, the proposed system can identify spots in environments that are inaccessible to the global positioning system. The proposed system has been evaluated by two experiments: image matching tests and a user study. The experimental results suggested the effectiveness of the system to help visually impaired individuals, including blind individuals, recall information about regularly-visited spots.

Keywords: spot reminder; visually impaired individuals; smartphone cameras; scale invariant feature transform; image matching; user study

1. Introduction

In 2014, the World Health Organization estimated the number of visually impaired individuals to be at approximately 285 million worldwide [1]. Many are trained by sighted assistants to move along their daily routes, for example, from home to the office. During such training, they are often taught information about important locations, called *spots*, along the routes. If they remember the information, their quality of life (QOL) would be maintained. Otherwise, they would be forced to suffer inconvenience. Figure 1 illustrates a typical situation in which the QOL of a visually impaired individual is strongly affected by whether the individual remembers the information about a spot. It is necessary to build an assistive system to help a visually impaired user recall information about environments.

A number of research groups have proposed obstacle detection systems to notify visually impaired users about the positions of obstacles which were detected using laser sensors [2–10], ultrasonic sensors [11–22] single charge-coupled devices (CCD) cameras [23–27], stereoscopic cameras [28–38], or RGB-D cameras [39–43]. These systems allow users to walk safely while avoiding obstacles (such as the pillar in Figure 1), even if they forget or do not know the positions of the obstacles.

The systems can warn users of obstacles in their vicinity but cannot tell the users what the objects are.

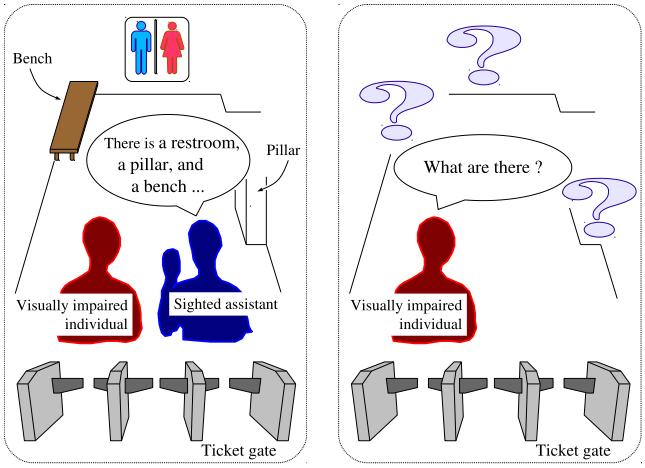


Figure 1. (**a**) One day; (**b**) Several days later. One day, a visually impaired individual visited a ticket gate in a station with a sighted assistant. The assistant taught the individual that there was a restroom in the front, a pillar on the right, and a bench on the left. Several days later, the individual returned alone, but forgot about the restroom and the bench. Therefore, the individual was not able to use them.

Other research groups have proposed assistive systems to recognize objects such as drug packages [44], podiums [45], classroom doors [45–48], and pathways [49,50], using barcodes [44,51], radio frequency identification tags [22,46,52,53], Bluetooth devices [54], augmented reality markers [45,47,48], circular markers [49], wireless network devices [50], or visible light communication devices [55,56]. These physical devices can help visually impaired users identify the objects, but in practice it is difficult to deploy them in an everyday environment.

Sensor-based systems have been developed to recognize color blocks [57], benches [58], tables [59], staircases [60–63], and elevators [64], using laser range sensors [62,63], laser pointers combined with a CCD camera [61], or Kinect sensors [57–59,64]. These systems allow visually impaired users to find and use target objects. For example, in Figure 1b, the individual can find the bench to take a rest. These systems can obtain information from the environment but cannot add information to the environment.

Smartphone-based systems have been proposed to navigate visually impaired individuals in indoor and outdoor environments. Elloumi et al. developed an algorithm for indoor pedestrian localization based on a smartphone camera fixed on a body harness [65]. Götzelmann et al. introduced an approach to combine a physical tactile map with an interactive application running on a smartphone for outdoor navigation [66]. These systems are also incapable of appending information to the environment.

Social platform systems have been proposed as a way of sharing barrier-free information among people with and without disabilities [67,68]. Anyone can upload information to maps on websites managed by the systems, and anyone can download the information from the websites. However, these systems require visually impaired users to search through a large amount of information uploaded by other people and to verify the downloaded information. Sekai Camera was a social platform system that was able to attach virtual information to the real world and display it on the screen of a smartphone. The Sekai Camera system had the same problem as the above platform systems and in addition was mainly designed for sighted people (The services of Sekai Camera were terminated in 2014.).

In this paper, we propose a smartphone-camera-based reminder system to help visually impaired users recall their own memories of spots that they visited. The memories are recorded as voice memos, which can be played back when the users return to the same spots. Spot-to-spot correspondence is determined by image matching between scene images obtained by the smartphone camera at the spots. The proposed system was implemented as an application on an Android smartphone and evaluated using two experiments: image matching tests and a user study.

This paper is organized as follows: Section 2 describes the outline of the spot reminder system, Section 3 explains the system implementation, Section 4 shows the experimental results, Section 5 discusses these results, and Section 6 concludes the paper.

2. Outline of the Spot Reminder System

Figure 2 illustrates the outline of the spot reminder system in two modes: record and playback.

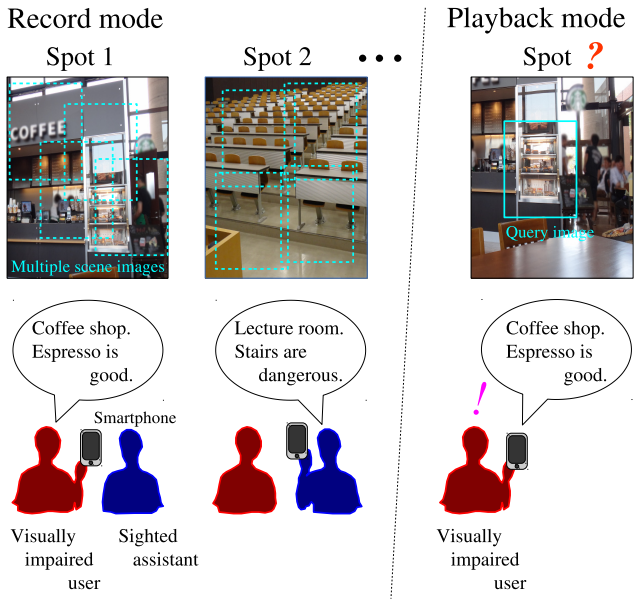


Figure 2. Outline of the spot reminder system.

In the record mode, at each spot, a visually impaired user or a sighted assistant takes multiple scene images (depicted by frames with dotted lines in Figure 2) using a smartphone camera and then records a voice memo about the spot (for example, "Coffee shop. Espresso is good.") on the smartphone. From the multiple images, feature points called *keypoints* are extracted using the Scale Invariant Feature Transform (SIFT) [69,70]. The images are then merged into one panoramic image using the SIFT-based image stitching technique [70,71]. The panoramic image and the voice memo are stored in a dictionary on the smartphone.

In the playback mode, when the user returns to one of the recorded spots, he or she takes a scene image for a search query (a frame with solid lines). SIFT-based image matching is performed between the query image and the dictionary images, and then the current spot is identified from the matching result. The smartphone plays the associated voice memo, which can help the user recall his or her memory related to the spot.

Figure 3 explains the reason why panoramic scene images are used in the proposed system. In Figure 3a, a visually impaired user attempts to identify the current spot on the basis of image matching between a single scene image taken in the record mode and a query image taken in the playback mode. These images are depicted by frames with dotted and solid lines, respectively. It is difficult for the visually impaired user to set a camera direction in the playback mode so that the frames overlap sufficiently, and therefore it is likely that the image matching would fail. Our previous

navigation system [72] and the VizMap localization system, proposed by Gleason et al. [73], used single scene images for spot identification and therefore had the same difficulty in image matching. In contrast, if a panoramic image is produced from multiple scene images with the help of a sighted assistant and is used for the image matching, it would be more promising as shown in Figure 3b. Image matching techniques were also used to recognize objects such as food packages in [74], where visually impaired users were required to set cameras toward target objects precisely. In our method, visually impaired users are less demanding.

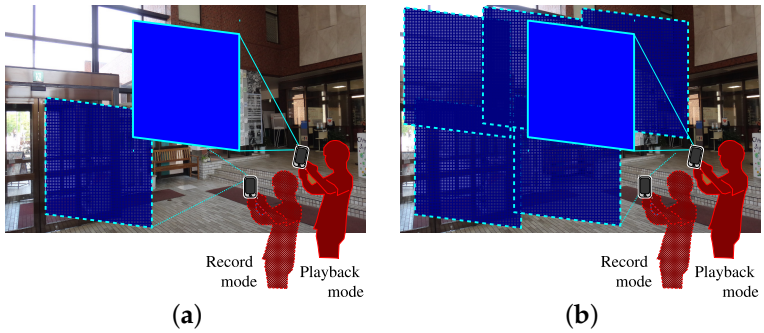


Figure 3. (**a**) Single scene image; (**b**) Panoramic scene image. If a single scene image is used for image matching, the system would fail to identify the current spot. By using a panoramic image, the current spot can be identified correctly.

2.1. Keypoint Extraction by SIFT

SIFT extracts keypoints through a detection phase followed by a description phase.

In the detection phase, a scene image is first enhanced by histogram equalization [75], which can increase the number of reliable keypoints [76]. The enhanced image is smoothed by the Gaussian filters of various scales (i.e., variances). By subtracting the adjacent Gaussian-smoothed images, Difference-of-Gaussian (DoG) images are produced. From the sequential DoG images, local extremum pixels are detected as keypoint candidates. Their sub-pixel positions and scales are obtained by interpolation based on the quadratic Taylor expansion of the DoG functions. By eliminating the low-contrast or on-the-edge candidates, the final keypoints are selected.

In the description phase, gradient magnitudes and orientations are computed for pixels in the Gaussian-smoothed images, and then the orientations of the keypoints are calculated from the 36-bin histograms of the gradient orientations weighted by their magnitudes. A squared region of interest (ROI) is set at each keypoint. Its size and orientation are determined from the scale and orientation of the keypoint, respectively. The ROI is divided into 4×4 blocks, and for each block, the 8-bin histogram of the gradient orientations weighted by the magnitudes is calculated. From the $4 \times 4 \times 8$ elements, a 128-dimensional feature vector is produced as the descriptor of the keypoint. The 128-dimensional feature vector is reasonably invariant against changes in scaling, rotation and illumination of images.

2.2. SIFT-Based Image Stitching in the Record Mode

Multiple scene images taken at each spot are sorted according to their timestamps, and are represented as I_n $(n = 1, \cdots, N)$. The first and second images, I_1 and I_2, are defined as reference and floating images, I^r and I^f, respectively. The following algorithm is iterated until all the images are processed.

1. From I^r and I^f, keypoints are extracted. They are represented as k_i^r $(i = 1, \cdots, I)$ and k_j^f $(j = 1, \cdots, J)$, respectively, and their 128-dimensional feature vectors are represented as $\boldsymbol{v}_i^r$ and $\boldsymbol{v}_j^f$, respectively. Each keypoint in the reference image is paired with the most similar keypoint in

the floating image. The similarity is evaluated by the following Euclidian distance between the 128-dimensional feature vectors of the keypoints:

$$d(i,j) = |\boldsymbol{v}_i^r - \boldsymbol{v}_j^f|^2. \quad (1)$$

The keypoint pairs are represented as p_k ($k = 1, \cdots, K$).

2. Some of the keypoint pairs come from the same objects observed in the reference and floating images. However, other pairs come from different objects and would cause errors in image stitching. They are removed by using the random sample consensus algorithm [77] as follows:

   (a) Four keypoint pairs, p_{k_1}, p_{k_2}, p_{k_3} and p_{k_4}, are chosen randomly.
   (b) A homography matrix [78], H, is calculated by applying the direct linear transformation algorithm [79] to the four keypoint pairs.
   (c) For the keypoint pair p_k, a back projection error, $err(p_k; H)$, is calculated. If $err(p_k; H) < \epsilon_H$, p_k is determined to be an *inlier*. Otherwise, *outlier*. The inliers and outliers represent the keypoint pairs of the same and different objects, respectively.
   (d) After iterating the above steps from (a) to (c), the algorithm determines the optimal homography matrix that produces the most inliers.

3. The floating image is transformed using the optimal homography matrix and merged into the reference image. The merged image is defined as a new reference image, and the next scene image, I_n ($n > 2$), is defined as a new floating image. The algorithm returns to step 1.

The final reference images are stored into the image dictionary on a smartphone.

2.3. SIFT-Based Image Matching in the Playback Mode

A query image is represented as I^q. The first dictionary image is selected as a checking image, I^c. The image matching is performed by the following algorithm:

1. Keypoints are extracted from I^q and I^c, and keypoint pairs are produced in the same manner as Section 2.2.
2. The geometrical relations between the keypoint pairs are evaluated on the basis of the following six criteria which have been proposed for pedestrian navigation [80,81]:

   (a) too few pairs
   (b) size consistency
   (c) direction consistency
   (d) two-dimensional affine constraint
   (e) area size
   (f) axis inversion

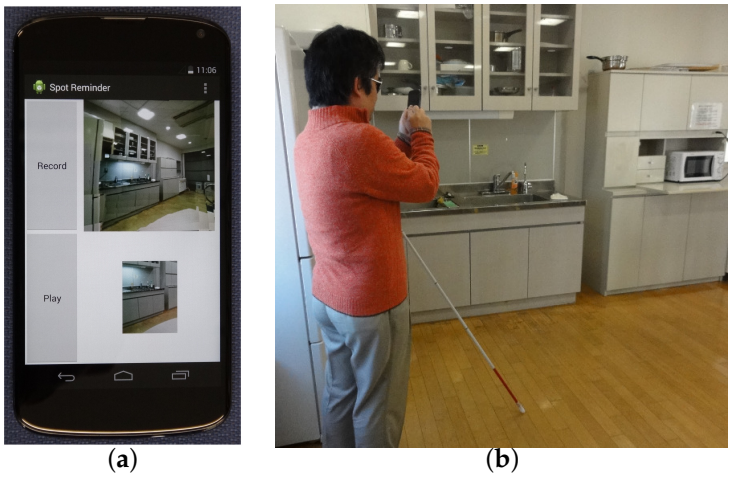
   If all the criteria are satisfied, I^c is determined to be matched with I^q and the algorithm is terminated.
3. The next dictionary image is selected as I^c, and the algorithm returns to step 1.

If I^q is not matched with any dictionary images, the algorithm determines that there are no match images.

3. System Implementation

The spot reminder system was implemented as an application on an Android Google Nexus 4 smartphone as shown in Figure 4a. The application displayed four components on the touch screen. The buttons "Record" and "Play" were able to activate the record and playback modes, respectively. Dictionary and query images allowed a sighted assistant to confirm the images. These images were able to be removed on demand and replaced by the activation buttons displayed full screen.

When a visually impaired user used the spot reminder system, the user stopped walking for safety and set the smartphone as shown in Figure 4b. In the record mode, the user or an assistant pushed the "Record" button to take multiple scene images and input a voice memo to the smartphone. In the playback mode, the user pushed the "Play" button to take a query image and heard the played-back voice memo.



(**a**) (**b**)

Figure 4. (**a**) Display of the system; (**b**) A user used the system. The spot reminder system.

4. Experiments

The proposed system was evaluated by image matching tests and a user study. The image matching tests evaluated the accuracy of the image matching between panoramic dictionary images and query images. The user study evaluated the effectiveness of the system in helping participants recall memories related to spots.

The parameters for keypoint extraction, image stitching and image matching were selected experimentally through preliminary experiments.

4.1. Image Matching Test

We carried out two types of image matching tests: identification and null response. In the identification test, the system was fed with query images taken at the same spots as dictionary images. The system was required to identify the dictionary images that corresponded to the query images. In the null-response test, the system was fed with query images taken at unknown spots. The system was required to reply that the queries were unknown.

4.1.1. Identification Test

Ten indoor spots and ten outdoor spots were selected for the test. At each spot, one panoramic image was produced, and four single images were taken. The twenty panoramic images were stored as dictionary images in a smartphone, and the eighty single images were input to the smartphone as queries.

Table 1 lists the accuracy of the image matching. *Correct identification* represents a case where the system correctly identified the dictionary image as matching the query image. *Incorrect identification* represents a case where the system mistakenly selected the dictionary image that did not correspond to the query image. *Null response* represents a case where the system determined that the dictionary included no images corresponding to the query image. In the identification test, the null response means a failure of image matching, because the dictionary included at least one image corresponding to each query image.

Table 1. Accuracy of the identification test.

	# of Images	Ratio
Correct identification	73	91%
Incorrect identification	0	0%
Null response	7	9%

In Table 1, seventy-three query images were identified correctly. No incorrect identifications were made, but in seven cases, the system failed to link the query images to those in the dictionary. Figures 5–8 show examples of matching results.

Figure 5 shows the result of image matching between a dictionary image (left) and a query image (right) taken at the same spot in a reinforced concrete building. In the figure, circles represent the scales of keypoints, segments in the circles represent their orientations, and lines between the keypoints represent keypoint pairs. Thirty-six keypoint pairs were obtained from the same objects at the spot. These keypoint pairs satisfied all the criteria of the playback mode, and the system correctly determined that these images were taken at the same spot.

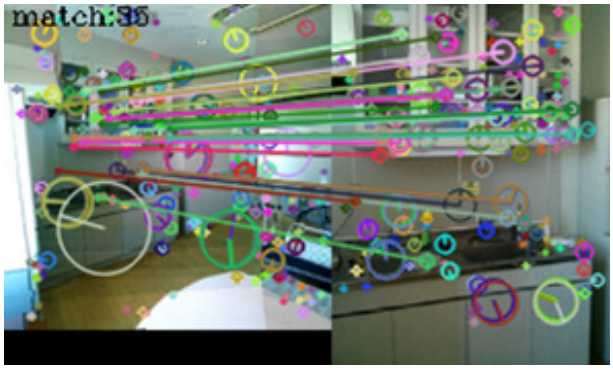


Figure 5. Result of image matching between dictionary and query images taken at the same indoor spot. The images have been resized for better visualization.

Figure 6 shows the result of image matching of images taken at the same outdoor spot. Thirty-seven keypoint pairs were obtained from the images, and the system identified the spot correctly.



Figure 6. Result of image matching between dictionary and query images taken at the same outdoor spot.

Figure 7 shows the result of image matching at outdoor and indoor spots. Nineteen keypoint pairs were obtained from the images but were generated between different objects. At the playback mode, the direction consistency criterion was not met. Consequently, the system correctly determined that the images were taken from different spots.

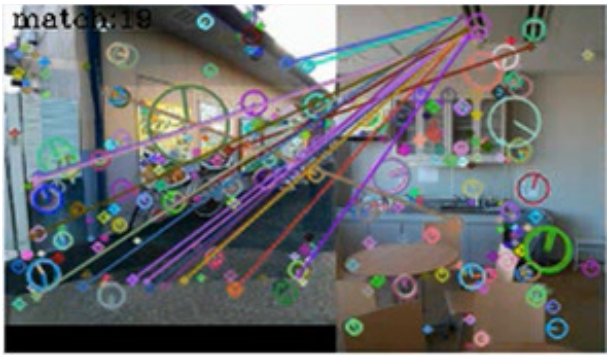


Figure 7. Result of image matching between dictionary and query images taken at different spots.

Figure 8 shows the result of image matching at the same outdoor spot. Twenty-four keypoint pairs were obtained, but many of them were concentrated in a small region. Consequently, the axis inversion criterion became unstable, and the system failed to identify the spot.



Figure 8. Result of image matching between dictionary and query images taken at the same outdoor spot.

4.1.2. Null-Response Test

Scene images were newly taken at twenty other spots and were input to the system as unknown queries. In the null-response test, null response means that the system succeeded in determining that the spots were unknown. All the query images were correctly determined to be unknown (Table 2).

Table 2. Accuracy of the null-response test.

	# of Images	Ratio
Incorrect identification	0	0%
Null response	20	100%

Figures 9 and 10 show the results of image matching between the dictionary images and the unknown query images. The system successfully determined that they were different spots.



Figure 9. Result of image matching between a dictionary image and an unknown query image.



Figure 10. Result of image matching between a dictionary image and an unknown query image.

4.2. User Study

We conducted a user study in which two blindfolded participants, graduate students in their twenties, attempted to recall their memories of five spots in a building located on their university campus (Figure 11).

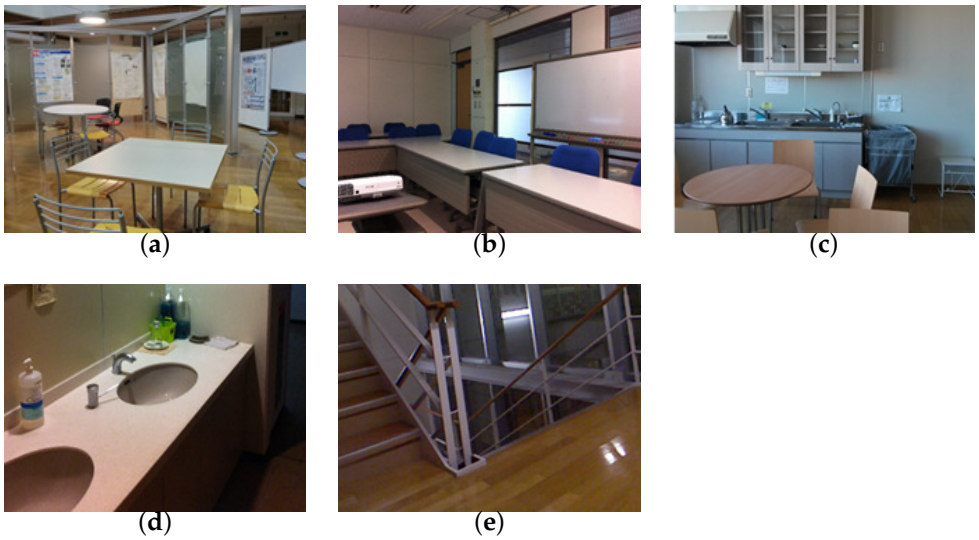


Figure 11. (**a**) Spot 1; (**b**) Spot 2; (**c**) Spot 3; (**d**) Spot 4; (**e**) Spot 5. Five spots used for the user study.

The participants were given fictitious information about these spots. For example, they were informed that the first spot was a convenience store including a restroom on the right side, an automated teller machine on the left side, a drink section at the right back, and a bread section at the left back. The reason why we used such fictitious information was that the participants were usually living in the building, and therefore they knew the true information about these spots. The fictitious information ensured that the participants did not to use their prior knowledge about these spots.

The user study was conducted as follows:

1. The participants were taken to the spots and given the fictitious information.
2. Ten minutes later, the participants were taken to the same spots again and asked to repeat the information without using the system.
3. The participants were then allowed to use the system and asked to repeat the information again.

Table 3 lists the results of the user study. In this table, "/" and "X" represent the cases where the participants reported the correct and incorrect information, respectively.

Table 3. Results of the user study.

		Participant 1		**Participant 2**	
		without System	**with System**	**without System**	**with System**
Spot	**Information**				
1	1	/	/	/	/
	2	/	/	/	/
	3	X	/	/	/
	4	/	/	/	/
2	1	/	/	/	/
	2	/	/	/	/
	3	/	/	/	/
	4	X	/	/	/
3	1	/	/	/	/
	2	/	/	X	/
	3	X	/	/	/
	4	/	/	/	/
4	1	X	/	/	/
	2	X	/	/	/
	3	/	/	/	/
	4	/	/	/	/
5	1	/	/	/	/
	2	X	/	/	/
	3	/	/	/	/
	4	/	/	X	/
Ratio		70%	100%	90%	100%

The first and second participants made six and two errors, respectively, when not using the proposed system. When using the system, they made no errors.

5. Discussion

The main contribution of the proposed system is to enable visually impaired users to associate arbitrary voice memos with arbitrary places and to automatically recover the voice memos when returning to the same places. No special devices or systems are required, except smartphones, and the users do not need to remember the places where the voice memos were recorded. Our preliminary investigation revealed that some visually impaired individuals had used conventional tape/IC recorders to record voice memos. These devices are incapable of linking voice memos with places,

requiring the users to search for voice memos corresponding to the current spots. The proposed system, in contrast, can identify the corresponding voice memos automatically, and therefore the users can obtain the necessary information more efficiently.

The advantage of the proposed system over the social platform systems is that visually impaired users can store only the desired information without having to discard irrelevant information uploaded by other people. For example, in the situation shown in Figure 2, the user may be interested only in the espresso coffee in the coffee shop. By excluding irrelevant detail in the record mode, the voice memo will only mention the espresso coffee in the playback mode. In a social platform system, if another user uploaded information about the other food and drink on sale, the user would need to search through the unnecessary information.

Many of the social platform systems [68] used the global positioning system (GPS) to determine the current positions of users. The GPS is unable to operate in many environments, such as the interiors of reinforced concrete buildings. The proposed system uses a vision-based positioning method previously used in other works [81,82]. This method can determine the current positions in locations where GPS is unable to operate.

In the user study, the participants were unable to recall at least 10% of the information about the spots without being prompted by the proposed system. When using the proposed system, they were able to recall all the information correctly. Although the participants were young graduate students and the intervals between the record and playback modes were only ten minutes, the results demonstrated that the proposed system was effective. This could be confirmed by user studies with older participants and longer time intervals.

Image matching methods based on single scene images [83,84] cannot be applied to our problem as described in Section 2. Therefore, we have improved the image matching method by combining with an image stitching method. Stitched images can cover the wider areas of scenes, and therefore visually impaired users can match the scenes even though the directions of smartphone cameras differ somewhat. The accuracy of image matching was not 100% as described in Section 4.1.1, and it would influence the spot identification. The image matching method should be improved. The fourth matching criterion (i.e., two-dimensional affine constraint) used in Section 2.3 is reasonably invariant against the changes of view points. The SIFT can absorb illumination changes to some extent, but cannot if the changes are too large. For example, the system was unable to match a dictionary image taken in sunlight to a query image of the same spot taken at sunset. This can be resolved by collecting scene images under a range of lighting conditions, but this is difficult to do in practice. New techniques, such as Colored SIFT [85], are needed to compensate the large illumination changes.

In the proposed method, visually impaired users should update spot information by themselves. As for the social platform systems, on the other hand, sighted volunteers can update the shared information via the network instead of visually impaired users. In this regard, the social platform systems are superior to the proposed system.

The proposed system can identify spots that a user visited once, but cannot recognize spots where he or she comes for the first time. In addition, if several spots have similar scenes, the system cannot identify the spots. We should improve the identification method and create a new recognition method for more convenience of the visually impaired.

The user study was carried out with a small number of blindfolded young people. With the experimental results, it cannot be said that the proposed system is effective for all visually impaired individuals. However, we think that the results would be able to suggest the effectiveness to people that became sightless lately. Especially, if they have the experiences of using smartphones, they can utilize our system easily because the proposed method is implemented as a typical application on a smartphone. In this paper, we mainly described the proposed system from the viewpoint of system development. In the future, we should evaluate the effectiveness of our system with actual visually impaired individuals.

The calculation time for image matching was less than one second when processing five spots. This would increase as more spots are added to the system. The image matching must be made more efficient to enable the system to deal with a larger number of spots.

6. Conclusions

The present paper proposed a spot reminder system to assist visually impaired users in recalling memories related to spots that they visited. The memories were recorded as voice memos, which were able to be played back when the user returned to the spots. The spot-to-spot correspondence was determined by SIFT-based image matching. The system was implemented as an application on an Android smartphone and evaluated by two experiments: image matching tests and a user study. The experimental results suggested the effectiveness of the system to help visually impaired individuals, including blind individuals, recall information about regularly-visited spots.

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

Author Contributions: H.T. initiated and led the study and was the main author of the manuscript. K.O. developed the algorithm, acquired the data, conducted experiments, and analyzed the results. M.A. was an advisor of the study. N.E. and S.M. participated in the design of the study. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. World Health Organization. World Health Organization, Media Centre, Visual Impairment And Blindness, Fact Sheet No. 282. 2014. Available online: http://www.who.int/mediacentre/factsheets/fs282/en/ (accessed on 1 August 2014).
2. Bolgiano, D.R.; Meeks, E. A Laser Cane for the Blind. *IEEE J. Quantum Electron.* **1967**, *3*, 268.
3. Benjamin, J.M.; Ali, N.A.; Schepis, A.F. A Laser Cane for the Blind. In Proceedings of the San Diego Biomedical Symposium, San Diego, CA, USA, 5–7 February 1973.
4. Benjamin, J.M., Jr. The Laser Cane. *J. Rehabil. Res. Dev.* **1974**, *BPR 10–22*, 443–450.
5. Engelbrektsson, P.; Karlsson, I.C.M.; Gallagher, B.; Hunter, H.; Petrie, H.; O'Neill, A.M. Developing a navigation aid for the frail and visually impaired. *Univers. Access Inf. Soc.* **2004**, *3*, 194–201.
6. Saegusa, S.; Yasuda, Y.; Uratani, Y.; Tanaka, E.; Makino, T.; Chang, J.Y. Development of a guide-dog robot: Human-robot interface considering walking conditions for a visually handicapped person. *Microsyst. Technol.* **2010**, *17*, 1169–1174.
7. Imadu, A.; Kawai, T.; Takada, Y.; Tajiri, T. Walking Guide Interface Mechanism and Navigation System for the Visually Impaired. In Proceedings of the 4th International Conference on Human System Interactions, Yokohama, Japan, 19–21 May 2011; pp. 34–39.
8. Gomez, J.V.; Sandnes, F.E. RoboGuideDog: Guiding Blind Users through Physical Environments with Laser Range Scanners. *Procedia Comput. Sci.* **2012**, *14*, 218–225.
9. Vera, P.; Zenteno, D.; Salas, J. A smartphone-based virtual white cane. *Pattern Anal. Appl.* **2014**, *17*, 623–632.
10. Lin, Q.; Han, Y. A Context-Aware-Based Audio Guidance System for Blind People Using a Multimodal Profile Model. *Sensors* **2014**, *14*, 18670–18700.
11. Pressey, N. Mowat Sensor. *Focus* **1977**, *11*, 35–39.
12. Morrissette, D.; Goodrich, G.; Hennessey, J. A follow-up study of the mowat sensor's applications, frequency of use, and maintenance reliability. *J. Vis. Impair. Blind.* **1981**, *75*, 244–247.
13. Okayasu, M. Newly developed walking apparatus for identification of obstructions by visually impaired people. *J. Mech. Sci. Technol.* **2010**, *24*, 1261–1264.
14. Wahab, M.H.A.; Talib, A.A.; Kadir, H.A.; Johari, A.; Noraziah, A.; Sidek, R.M.; Mutalib, A.A. Smart Cane: Assistive Cane for Visually-impaired People. *Int. J. Comput. Sci. Issues* **2011**, *8*, 21–27.

15. Dambhare, S.; Sakhare, A. Smart stick for Blind: Obstacle Detection, Artificial vision and Real-time assistance via GPS. In Proceedings of the 2nd National Conference on Information and Communication Technology, Chennai, India, 23–24 December 2011; Foundation of Computer Science: New York, NY, USA, 2011; pp. 31–33.
16. Shoval, S.; Borenstein, J.; Koren, Y. The NavBelt—A Computerized Travel Aid for the Blind Based on Mobile Robotics Technology. *IEEE Trans. Biomed. Eng.* **1998**, *45*, 1376–1386.
17. Bharathi, S.; Ramesh, A.; Vivek, S.; Kumar, J. Effective Navigation for Visually Impaired by Wearable Obstacle Avoidance System. *Int. J. Power Control Signal Comput.* **2012**, *3*, 51–53.
18. Cardin, S.; Thalmann, D.; Vexo, F. A wearable system for mobility improvement of visually impaired people. *Vis. Comput.* **2007**, *23*, 109–118.
19. Bahadir, S.K.; Koncar, V.; Kalaoglu, F. Wearable obstacle detection system fully integrated to textile structures for visually impaired people. *Sens. Actuators A Phys.* **2012**, *179*, 297–311.
20. Ulrich, I.; Borenstein, J. The GuideCane—Applying Mobile Robot Technologies to Assist the Visually Impaired. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* **2001**, *31*, 131–136.
21. Mahmud, M.H.; Saha, R.; Islam, S. Smart walking stick—An electronic approach to assist visually disabled persons. *Int. J. Sci. Eng. Res.* **2013**, *4*, 111–114.
22. Rohan, P.; Ankush, G.; Vaibhav, S.; Dheeraj, M.; Balakrishnan, M.; Kolin, P.; Dipendra, M. Smart cane for the visually impaired: Technological solutions for detecting knee-above obstacles and accessing public buses. In Proceedings of the 11th International Conference on Mobility and Transport for Elderly and Disabled Persons, Montreal, QC, Canada, 18–21 June 2007.
23. Peng, E.; Peursum, P.; Li, L.; Venkatesh, S. A Smartphone-Based Obstacle Sensor for the Visually Impaired. In *Ubiquitous Intelligence and Computing*; Lecture Notes in Computer Science; Yu, Z., Liscano, R., Chen, G., Zhang, D., Zhou, X., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6406, pp. 590–604.
24. Tapu, R.; Mocanu, B.; Bursuc, A.; Zaharia, T. A Smartphone-Based Obstacle Detection and Classification System for Assisting Visually Impaired People. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Sydney, Australia, 1–8 December 2013; pp. 444–451.
25. Kishino, T.; Zhe, S.; Ruggero, M. A Fast and Precise HOG-Adaboost Based Visual Support System Capable to Recognize Pedestrian and Estimate Their Distance. In Proceedings of the International Conference on Image Analysis and Processing Workshops, Naples, Italy, 9–13 September 2013; pp. 20–29.
26. Praveen, R.G.; Paily, R.P. Blind Navigation Assistance for Visually Impaired based on Local Depth Hypothesis from a Single Image. *Procedia Eng.* **2013**, *64*, 351–360.
27. Caldini, A.; Fanfani, M.; Colombo, C. Smartphone-Based Obstacle Detection for the Visually Impaired. In *Image Analysis and Processing—ICIAP 2015: 18th International Conference, Genoa, Italy, September 7–11, 2015, Proceedings, Part I*; Springer: Cham, Switzerland, 2015; pp. 480–488.
28. 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, Quebec City, QC, Canada, 1–15 August 2002; Volume 3, pp. 974–977.
29. Balakrishnan, G.; Sainarayanan, G.; Nagarajan, R.; Yaacob, S. A Stereo Image Processing System for Visually Impaired. *World Acad. Sci. Eng. Technol.* **2006**, *20*, 206–215.
30. Balakrishnan, G.; Sainarayanan, G.; Nagarajan, R.; Yaacob, S. Wearable Real-Time Stereo Vision for the Visually Impaired. *Eng. Lett.* **2007**, *14*, 1–9.
31. Dunai, L.; Fajarnes, G.P.; Praderas, V.S.; Garcia, B.D.; Lengua, I.L. Real-Time Assistance Prototype—A New Navigation Aid for Blind People. In Proceedings of the 36th Annual Conference on IEEE Industrial Electronics Society, Glendale, CA, USA, 7–10 November 2010; pp. 1173–1178.
32. Velzquez, R.; Maingreaud, F.; Pissaloux, E.E. Intelligent Glasses: A New Man-Machine Interface Concept Integrating Computer Vision and Human Tactile Perception. In Proceedings of the EuroHaptics 2003, Dublin, Ireland, 6–9 July 2003; pp. 456–460.
33. Meers, S.; Ward, K. Substitute three-dimensional perception using depth and colour sensors. In Proceedings of the 2007 Australasian Conference on Robotics and Automation, Brisbane, Australia, 10–12 December 2007; pp. 1–5.
34. Kim, D.; Kim, K.; Lee, S. Stereo Camera Based Virtual Cane System with Identifiable Distance Tactile Feedback for the Blind. *Sensors* **2014**, *14*, 10412–10431.

35. Molton, N.; Se, S.; Brady, J.; Lee, D.; Probert, P. A stereo vision-based aid for the visually impaired. *Image Vis. Comput.* **1998**, *16*, 251–263.
36. Ikarashi, M.; Yokote, H.; Takizawa, H.; Yamamoto, S. Walking support system using stereo data for blind person. In Proceedings of the IEICE General Conference, Hiroshima, Japan, 28–31 March 2000; Volume 2, p. 337.
37. Saito, T.; Takizawa, H.; Yamamoto, S. A Display System of Obstacle Positions for Visible Disabled Persons. In Proceedings of the IEICE General Conference, Tokyo, Japan, 27–30 March 2002; Volume 2, p. 316.
38. Zelek, J.; Audette, R.; Balthazaar, J.; Dunk, C. *A Stereo-Vision System for the Visually Impaired*; Technical Report; University of Guelph: Guelph, ON, Canada, 2000.
39. Khan, A.; Moideen, F.; Lopez, J.; Khoo, W.L.; Zhu, Z. KinDectect: Kinect Detecting Objects. In Proceedings of the 13th International Conference on Computers Helping People with Special Needs, Linz, Austria, 11–13 July 2012; pp. 588–595.
40. Bernabei, D.; Ganovelli, F.; Benedetto, M.D.; Dellepiane, M.; Scopigno, R. A Low-Cost Time-Critical Obstacle Avoidance System for the Visually Impaired. In Proceedings of the International Conference on Indoor Positioning and Indoor Navigation, Guimarães, Portugal, 21–23 September 2011; pp. 21–23.
41. Salerno, M.; Re, M.; Cristini, A.; Susi, G.; Bertola, M.; Daddario, E.; Capobianco, F. AudiNect: An Aid for the Autonomous Navigation of Visually Impaired People Based on Virtual Interface. *Int. J. Hum. Comput. Interact.* **2013**, *4*, 25–33.
42. Lee, Y.H.; Medioni, G. RGB-D camera Based Navigation for the Visually Impaired. In Proceedings of the RSS 2011 RGB-D: Advanced Reasoning with Depth Camera Workshop, Los Angeles, CA, USA, 27 June 2011; pp. 1–6.
43. Orita, K.; Takizawa, H.; Aoyagi, M.; Ezaki, N.; Shinji, M. Obstacle Detection by the Kinect Cane System for the Visually Impaired. In Proceedings of the 2013 IEEE/SICE International Symposium on System Integration, Kobe, Japan, 15–17 December 2013; Volume 1, pp. 115–118.
44. Lee, H.P.; Sheu, T.F. Building a Portable Talking Medicine Reminder for Visually Impaired Persons. In Proceedings of the Sixth International Conference on Future Computational Technologies and Applications, Venice, Italy, 25–29 May 2014; pp. 13–14.
45. Manduchi, R.; Coughlan, J.; Ivanchenko, V. Search Strategies of Visually Impaired Persons Using a Camera Phone Wayfinding System. In Proceedings of the Computers Helping People with Special Needs, Linz, Austria, 9–11 July 2008; pp. 1135–1140.
46. Kulyukin, V.; Gharpure, C.; Nicholson, J. RFID in Robot-Assisted Indoor Navigation for the Visually Impaired. In Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan, 28 September–2 October 2004; pp. 1979–1984.
47. Zöllner, M.; Huber, S.; Jetter, H.C.; Reiterer, H. NAVI—A Proof-of-Concept of a Mobile Navigational Aid for Visually Impaired Based on the Microsoft Kinect. In Proceedings of the 13th IFIP TC13 Conference on Human-Computer Interaction, Lisbon, Portugal, 5–9 September 2011; pp. 584–587.
48. Halabi, O.; Al-Ansari, M.; Halwani, Y.; Al-Mesaifri, F.; Al-Shaabi, R. Navigation Aid for Blind People Using Depth Information and Augmented Reality Technology. In Proceedings of the NICOGRAPH International 2012, Bali, Indonesia, 2–3 July 2012; pp. 120–125.
49. Fernandes, H.; Costa, P.; Filipe, V.M.; Hadjileontiadis, L.; Barroso, J. Stereo vision in blind navigation assistance. In Proceedings of the World Automation Congress, Kobe, Japan, 19–23 September 2010.
50. Au, A.W.S.; Feng, C.; Valaee, S.; Reyes, S.; Sorour, S.; Markowitz, S.N.; Gold, D.; Gordon, K.; Eizenman, M. Indoor Tracking and Navigation Using Received Signal Strength and Compressive Sensing on a Mobile Device. *IEEE Trans. Mob. Comput.* **2013**, *12*, 2050–2062.
51. Kulyukin, V.; Nicholson, J.; Coster, D. Shoptalk: Toward Independent Shopping by People with Visual Impairments. In Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility, Halifax, NS, Canada, 13–15 October 2008; ACM: New York, NY, USA, 2008; pp. 241–242.
52. Debnath, N.; Hailani, Z.A.; Jamaludin, S.; Aljunid, I.D.S.A.K. An Electronically Guided Walking Stick for the Blind. In Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Istanbul, Turkey, 25–28 October 2001; Volume 2, pp. 1377–1379.
53. Tatsumi, H.; Murai, Y.; Miyakawa, M. RFID for aiding the visually impaired recognize surroundings. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Montreal, QC, Canada, 7–10 October 2007; pp. 3719–3724.

54. Markiewicz, M.; Skomorowski, M. Public Transport Information System for Visually Impaired and Blind People. In *Transport Systems Telematics*; Communications in Computer and Information Science; Mikulski, J., Ed.; Springer: Berlin/Heidelberg, Germany, 2010; Volume 104, pp. 271–277.
55. Nakajima, M.; Haruyama, S. New indoor navigation system for visually impaired people using visible light communication. *EURASIP J. Wirel. Commun. Netw.* **2013**, *2013*, 1–10.
56. Nakazawa, Y.; Makino, H.; Nishimori, K.; Wakatsuki, D.; Komagata, H. LED-Tracking and ID-Estimation for Indoor Positioning Using Visible Light Communication. In Proceedings of the Fifth International Conference on Indoor Positioning and Indoor Navigation, Busan, Korea, 27–30 October 2014; pp. 1–8.
57. Gomez, J.D.; Mohammed, S.; Bologna, G.; Pun, T. Toward 3D scene understanding via audio-description: Kinect-iPad fusion for the visually impaired. In Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility, Dundee, UK, 26–28 October 2011; ACM: New York, NY, USA, 2011; pp. 293–294.
58. Takizawa, H.; Yamaguchi, S.; Aoyagi, M.; Ezaki, N.; Mizuno, S. Kinect cane: An assistive system for the visually impaired based on the concept of object recognition aid. *Pers. Ubiquitous Comput.* **2015**, *19*, 955–965.
59. Wang, Z.; Liu, H.; Wang, X.; Qian, Y. Segment and Label Indoor Scene Based on RGB-D for the Visually Impaired. In *MultiMedia Modeling*; Lecture Notes in Computer Science; Gurrin, C., Hopfgartner, F., Hurst, W., Johansen, H., Lee, H., O'Connor, N., Eds.; Springer: Heidelberg, Germany, 2014; Volume 8325, pp. 449–460.
60. Filipe, V.; Fernandes, F.; Fernandes, H.; Sousa, A.; Paredes, H.; Barroso, J. Blind Navigation Support System based on Microsoft Kinect. In Proceedings of the 4th International Conference on Software Development for Enhancing Accessibility and Fighting Info-Exclusion (DSAI 2012), Douro Region, Portugal, 19–22 July 2012; pp. 94–101.
61. Yasumuro, Y.; Murakami, M.; Imura, M.; Kuroda, T.; Manabe, Y.; Chihara, K. E-cane with situation presumption for the visually impaired. In Proceedings of the 7th ERCIM International Workshop on User Interfaces for All, Paris, France, 24–25 October 2002; Springer: Berlin/Heidelberg, Germany, 2003; pp. 409–421.
62. Ueda, T.; Kawata, H.; Tomizawa, T.; Ohya, A.; Yuta, S. Visual Information Assist System Using 3D SOKUIKI Sensor for Blind People, System Concept and Object Detecting Experiments. In Proceedings of the 32nd Annual Conference on IEEE Industrial Electronics, Paris, France, 7–10 November 2006; pp. 3058–3063.
63. Ishiwata, K.; Sekiguchi, M.; Fuchida, M.; Nakamura, A. Basic study on step detection system for the visually impaired. In Proceedings of the 2013 IEEE International Conference on Mechatronics and Automation, Takamatsu, Japan, 4–7 August 2013; pp. 1332–1337.
64. Kuramochi, Y.; Takizawa, H.; Aoyagi, M.; Ezaki, N.; Shinji, M. Recognition of Elevators with the Kinect Cane System for the Visually Impaired. In Proceedings of the 2014 IEEE/SICE International Symposium on System Integration, Tokyo, Japan, 13–15 December 2014; Volume 1, pp. 128–131.
65. Elloumi, W.; Guissous, K.; Chetouani, A.; Canals, R.; Leconge, R.; Emile, B.; Treuillet, S. Indoor navigation assistance with a Smartphone camera based on vanishing points. In Proceedings of the International Conference on Indoor Positioning and Indoor Navigation, Montbéliard-Belfort, France, 28–31 October 2013; pp. 1–9.
66. Götzelmann, T.; Winkler, K. SmartTactMaps: A Smartphone-based Approach to Support Blind Persons in Exploring Tactile Maps. In Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments, Corfu, Greece, 1–3 July 2015; ACM: New York, NY, USA, 2015; doi:10.1145/2769493.2769497.
67. Tuomisto, J.; Rajamäki, J. YPOP Indoor Navigation and Service Information System for Public Environments. In Proceedings of the 11th WSEAS International Conference on Communications, Crete Island, Greece, 26–28 July 2007; World Scientific and Engineering Academy and Society (WSEAS): Stevens Point, WI, USA, 2007; Volume 11, pp. 208–213.
68. Miura, T.; Yabu, K.I.; Sakajiri, M.; Ueda, M.; Suzuki, J.; Hiyama, A.; Hirose, M.; Ifukube, T. Social Platform for Sharing Accessibility Information Among People with Disabilities: Evaluation of a Field Assessment. In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, Bellevue, WA, USA, 21–23 October 2013; ACM: New York, NY, USA, 2013; p. 65.
69. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.
70. OpenCV. 2006. Available online: http://opencv.org (accessed on 23 December 2016).

71. Brown, M.; Lowe, D. Automatic Panoramic Image Stitching Using Invariant Features. *Int. J. Comput. Vis.* **2007**, *74*, 59–73.
72. Takizawa, H.; Orita, K.; Aoyagi, M.; Ezaki, N.; Mizuno, S. A Spot Navigation System for the Visually Impaired by Use of SIFT-Based Image Matching. In Proceedings of the 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, 2–7 August 2015; Volume 9178, pp. 160–167.
73. Gleason, C.; Guo, A.; Laput, G.; Kitani, K.; Bigham, J.P. VizMap: Accessible Visual Information through Crowdsourced Map Reconstruction. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility, Sopot, Poland, 6–8 June 2013; ACM: New York, NY, USA, 2016; pp. 273–274.
74. Matusiak, K.; Skulimowski, P.; Strumiłło, P. Object recognition in a mobile phone application for visually impaired users. In Proceedings of the IEEE 6th International Conference on Human System Interaction, Sopot, Poland, 6–8 June 2013; pp. 1–6.
75. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 3rd ed.; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2008.
76. Orita, K.; Takizawa, H.; Aoyagi, M.; Ezaki, N.; Shinji, M. *Basic Study on Memory Retrieval Assisting for the Visually Impaired by Use of SIFT-Based Image Matching*; IEICE Technical Report (PRMU2014-94); The Institute of Electronics, Information and Communication Engineers: Tokyo, Japan, 2015; Volume 114, pp. 93–98.
77. Fischler, M.A.; Bolles, R.C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* **1981**, *24*, 381–395.
78. Richard, S. *Computer Vision—Algorithms and Applications*; Springer: London, UK, 2011.
79. Solem, J.E. *Programming Computer Vision with Python—Tools and Algorithms for Analyzing Images*; O'Reilly Media Inc.: Sebastopol, CA, USA, 2012.
80. Kameda, Y.; Ohta, Y. Image Retrieval of First-Person Vision for Pedestrian Navigation in Urban Area. In Proceedings of the 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 364–367.
81. Kurata, T.; Kourogi, M.; Ishikawa, T.; Kameda, Y.; Aoki, K.; Ishikawa, J. Indoor-Outdoor Navigation System for Visually-Impaired Pedestrians: Preliminary Evaluation of Position Measurement and Obstacle Display. In Proceedings of the 15th IEEE International Symposium on Wearable Computers, San Francisco, CA, USA, 12–15 June 2011; pp. 123–124.
82. Treuillet, S.; Royer, E.; Chateau, T.; Dhome, M.; Lavest, J.M. Body Mounted Vision System for Visually Impaired Outdoor and Indoor Wayfinding Assistance. In Proceedings of the Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Assistive Technology for All Ages, Granada, Spain, 28–31 August 2007; Hersh, M., Ed.; 2007; pp. 1–6.
83. Lowe, D.G. Object recognition from local scale-invariant features. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; Volume 2, pp. 1150–1157.
84. Karami, E.; Prasad, S.; Shehata, M. Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images. In Proceedings of the 2015 Newfoundland Electrical and Computer Engineering Conference, New York, NY, USA, 5–6 November 2015; pp. 1–4.
85. Abdel-Hakim, A.E.; Farag, A.A. CSIFT: A SIFT descriptor with color invariant characteristics. In Proceedings of IEEE Conference on ComputerVision and Pattern Recognition, New York, NY, USA, 17–22 June 2006; pp. 1978–1983.



© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).