Article

Smartphone-Based Escalator Recognition for the Visually Impaired

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

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

Abstract: It is difficult for visually impaired individuals to recognize escalators in everyday environments. If the individuals ride on escalators in the wrong direction, they will stumble on the steps. This paper proposes a novel method to assist visually impaired individuals in finding available escalators by the use of smartphone cameras. Escalators are recognized by analyzing optical flows in video frames captured by the cameras, and auditory feedback is provided to the individuals. The proposed method was implemented on an Android smartphone and applied to actual escalator scenes. The experimental results demonstrate that the proposed method is promising for helping visually impaired individuals use escalators.

1. Introduction

In 2014, the World Health Organization reported that the number of visually impaired individuals was estimated to be approximately 285 million worldwide [1]. Many of them use white canes to detect obstacles ahead of them, but the detection ranges are short. Guide dogs are also used for navigation, but they need long training periods and large budgets. Therefore, it is necessary to build assistive systems [2] to help the visually impaired.

Several research groups have proposed cane-type systems [3–12], belt-type systems [13–16], helmet-type systems [17–20], wearable systems [21–23], glasses-type systems [24,25], and robot systems [26–29] to detect obstacles and recognize objects in environments. These assistive systems were handmade; therefore, it is difficult for visually impaired individuals to obtain them.

Other research groups have proposed assistive systems based on general smartphones. Obstacle detection systems were proposed in [14,26,30]. Dumitras et al. [31] proposed a mobile text-recognition system to allow the visually impaired to access text information. Tekin et al. [32] developed a system to detect and read LED/LCD digit characters of a certain font. Zhang et al. [33] proposed a mobile recognition system of braille characters [34] on public telephones or guide plates. Sara et al. [35] built a color recognition system for clothing coordination based on HSL color space processing. Matusiak et al. [36] proposed a recognition system of food or medicine packages based on Scale-Invariant Feature Transform (SIFT) [37] or Features from accelerated segment test (FAST) [38]. Ivanchenko et al. proposed a mobile phone system to allow the visually impaired to know the positions of crosswalks [39]. They also proposed a walk light detection system to let visually impaired individuals know when it is time to cross [40]. These systems can recognize static objects around

visually impaired individuals. However, in real environments, there are many dynamic objects such as people and cars.

Tapu et al. [41–43] proposed categorization methods of dynamic objects such as cars, bicycles, and pedestrians, as well as static obstructions based on computer vision techniques. The methods were implemented on portable systems composed of smartphones and several devices mounted on chest harnesses. These systems can notify visually impaired individuals about dynamic objects, but cannot help the individuals use the objects. The systems can only warn the individuals not to collide with the objects. In daily life, however, even visually impaired individuals often need to use dynamic objects, such as moving walkways and rotating doors.

In this paper, we focus on escalators. In general, visually impaired individuals estimate the positions of escalators based on motor sounds and then walk to the estimated positions. Subsequently, they grope for the belts of the escalators, and confirm their movement directions. If the directions are suitable, they can ride on the escalators. In actual escalator scenes, however, it is difficult to find the escalator belts; therefore, the individuals often fail to determine the movement directions. If the escalators move in the wrong directions, it can be dangerous for the visually impaired.

This paper proposes an escalator recognition method for visually impaired individuals. This method can detect the positions of escalators and determine their movement directions from videos obtained with a smartphone camera. The method can also provide auditory feedback to let the individuals know the recognition results. The proposed method is implemented on an Android smartphone.

Section 2 describes the outline of the proposed method, Section 3 shows experimental results from actual escalator scenes, Section 4 discusses the proposed method, and Section 5 concludes the paper.

2. Outline of the Proposed Method

Figure 1 shows the outline of the proposed method. First, when a visually impaired user predicts that he or she is in front of an escalator, the user sets his or her smartphone vertically and takes a video of the scene with the camera, as shown in Figure 2. The user can pan the camera to search for an escalator, if necessary. The video is divided into frames, from which corner points are detected. Optical flows are computed at the corners in two successive frames. A homography matrix $\mathbf{H}$ is estimated by applying the random sample consensus (RANSAC) algorithm [44] to the optical flows, which are classified into two categories: *inliers* and *outliers*. The inlier optical flows come from the camera motion, because it affects the entire image. The camera motion is obtained by averaging the inlier optical flows. The frame at $t = t$ is transformed into the frame at $t = t + 1$ by using an image registration technique based on the Homography matrix. A difference image is made from the transformed frame at $t = t$ and the frame at $t = t + 1$. In the difference image, moving objects appear as regions with high intensities. These regions are extracted as masks by a binarization operator followed by morphological operations [45] for shape smoothing. The camera motion is subtracted from the optical flows on the masks in the frame at $t = t + 1$. The final optical flows represent the direction of the moving objects (i.e., steps). Depending on the number and direction of the optical flows, the system recognizes the escalator and informs the user.

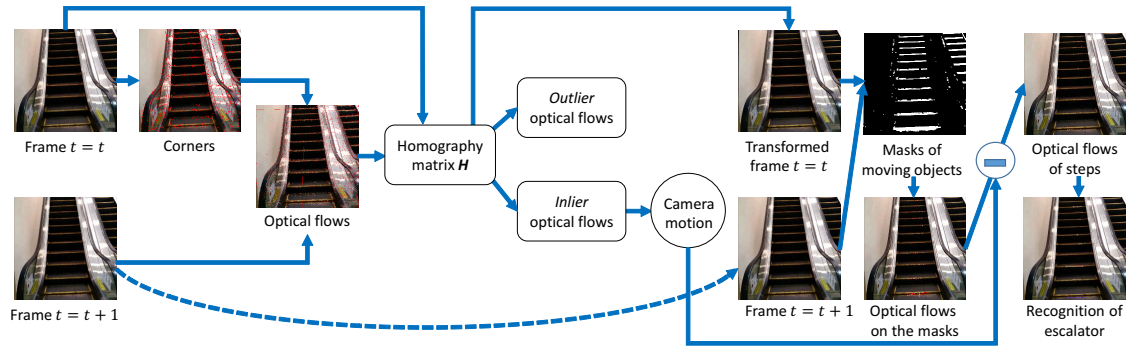The method is described in detail in the following sections.
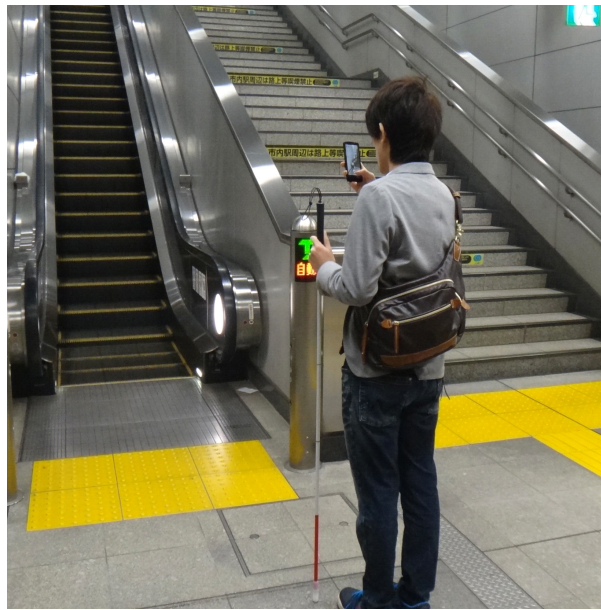
Figure 1. Outline of the proposed method.

Figure 2. A user takes a video with a smartphone camera.

2.1. Corner Detection

Escalator steps have concave–convex surfaces for skidproof purposes, and are highlighted with an accent color such as yellow. Therefore, they are often observed as a set of corner points with strong contrast in video frames. Such corner points are detected as described below.

Let $I(x,y,t)$ denote the intensity of a pixel $p(x,y)$ in a frame at a certain time t, and λ_p denote the minimal eigenvalue of the following matrix [46]:

$$\mathbf{M} = \begin{bmatrix} \sum\limits_{S_p} \left(\frac{dI}{dx}\right)^2 & \sum\limits_{S_p} \left(\frac{dI}{dx}\frac{dI}{dy}\right) \\ \sum\limits_{S_p} \left(\frac{dI}{dx}\frac{dI}{dy}\right) & \sum\limits_{S_p} \left(\frac{dI}{dy}\right)^2 \end{bmatrix}, \tag{1}$$

where S_p represents a small region with its center at pixel p in the frame. The minimal eigenvalues are calculated at all the pixels, and if a minimal eigenvalue λ_p is the maximum in S_p, it is eliminated. Among the remaining eigenvalues, the maximum eigenvalue λ_{max} is determined, and eigenvalues less than $Q_\lambda\%$ of λ_{max} are also eliminated. The remaining eigenvalues are denoted by $\lambda_{(1)}, \lambda_{(2)}, \cdots$ ($\lambda_{(1)} \geq \lambda_{(2)} \geq \cdots$), where $\lambda_{(1)}$ is equivalent to λ_{max}. The pixel at λ_{max} is extracted as the first corner point. If the distance between $\lambda_{(i)}$ and $\lambda_{(j)}$ ($i > j$) is larger than a predefined threshold d_{cd}, the pixel at

$\lambda_{(i)}$ is also extracted as a corner point. In this way, N_{fn} corner points are extracted. The corner points can be used as clues for the recognition of the escalator steps.

2.2. Optical Flow Computation

In order to recognize the movement direction of the escalator steps, we used the gradient-based optical flow detection method, where the intensities of corresponding pixels are assumed to remain unchanged in two successive frames. The assumption is represented by

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t), \tag{2}$$

where $(\delta x, \delta y)$ is the displacement of a pixel $p(x, y)$ during an interval time δt. By applying the Taylor-expansion to the right term of Equation (2), we can obtain

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + \frac{\partial I}{\partial x}\delta x + \frac{\partial I}{\partial y}\delta y + \frac{\partial I}{\partial t}\delta t + O_I, \tag{3}$$

where O_I is a high-order term, which can be omitted. From Equations (2) and (3), the following equation is obtained:

$$\frac{\partial I}{\partial x}\frac{\delta x}{\delta t} + \frac{\partial I}{\partial y}\frac{\delta y}{\delta t} + \frac{\partial I}{\partial t} = 0. \tag{4}$$

Let u_x and u_y be the x and y components of the optical flow at $p(x, y)$, respectively, and I_x, I_y, and I_t be the derivatives of $I(x, y, t)$ in the corresponding directions. By using u_x, u_y, I_x, I_y, and I_t, Equation (4) is converted to

$$I_x u_x + I_y u_y + I_t = 0. \tag{5}$$

Equation (5) is known as the optical flow constraint equation, which has two variables: u_x and u_y. This equation can be solved with the Lucas–Kanade algorithm (LKA) [47], which assumes that optical flows are uniform in local regions. Let us consider a small region whose center is at $p(x, y)$. The region size is set to be $N_{rs} = M_{rs} \times M_{rs}$ pixels. The assumption gives us the following N_{rs} equations:

$$\begin{cases} I_x^{(1)} u_x + I_y^{(1)} u_y = -I_t^{(1)}, \\ I_x^{(2)} u_x + I_y^{(2)} u_y = -I_t^{(2)}, \\ \quad\vdots \\ I_x^{(N_{rs})} u_x + I_y^{(N_{rs})} u_y = -I_t^{(N_{rs})}. \end{cases} \tag{6}$$

These equations are rewritten as follows:

$$\mathbf{Au} = -\mathbf{b}, \tag{7}$$

where

$$\mathbf{A} = \begin{bmatrix} I_x^{(1)} & I_y^{(1)} \\ \vdots & \vdots \\ I_x^{(N_{rs})} & I_y^{(N_{rs})} \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_x \\ u_y \end{bmatrix}, \mathbf{b} = \begin{bmatrix} I_t^{(1)} \\ \vdots \\ I_t^{(N_{rs})} \end{bmatrix}. \tag{8}$$

The optical flow $\mathbf{u}$ can be obtained by applying the least squares method to Equation (7). The LKA is sensitive to noise in the frames; therefore, we used the extended LKA [48] based on the pyramidal multiresolution analysis, where optical flows in a frame at a resolution are computed from those in another frame at a lower resolution.

2.3. Homography Transformation for Image Registration

Frames are often deformed due to accidental movement of the user's hands when taking the video. The deformation is compensated by the homography transformation [49], which is a kind of planar projective transformation. Let $(x_i, y_i, w_i)^T$ and $(x_i', y_i', w_i')^T$ denote the 2D homogeneous coordinates of the start and end points of the *i*-th optical flow, respectively $(i = 1, 2, \cdots, N_{fn})$. The homography transformation is represented by

$$
\begin{bmatrix} x_i' \\ y_i' \\ w_i' \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ w_i \end{bmatrix},
\tag{9}
$$

where

$$
\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix},
\tag{10}
$$

is known as the homography matrix and is computed using the direct linear transformation (DLT) algorithm [49].

2.3.1. DLT Algorithm

In the DLT algorithm, the homography transformation is represented by

$$
\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1 x_1' & -y_1 x_1' & -x_1' \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1 y_1' & -y_1 y_1' & -y_1' \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2 x_2' & -y_2 x_2' & -x_2' \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2 y_2' & -y_2 y_2' & -y_2' \\ & \vdots & & & \vdots & & & \vdots & \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix} = 0,
\tag{11}
$$

and it is rewritten as

$$
\mathbf{Ah} = 0,
\tag{12}
$$

where $\mathbf{A}$ is a $2N_{fn} \times 9$ matrix. The parameter vector $\mathbf{h}$ is another expression of the homography matrix, and can be obtained by the singular value decomposition (SVD) [50], which converts the matrix as follows:

$$
\mathbf{A} = \mathbf{UDV}^T,
\tag{13}
$$

where $\mathbf{U}$ is a $2N_{fn} \times 9$ orthogonal matrix, $\mathbf{D}$ is a 9×9 diagonal matrix, and $\mathbf{V}$ is a 9×9 orthogonal matrix. Each diagonal element $d_i(d_1 \geq d_2 \geq \cdots \geq d_i \geq \cdots \geq d_9 \geq 0)$ of $\mathbf{D}$ is a singular value of $\mathbf{A}$, and also the square root of the eigenvalue of $\mathbf{A}^T\mathbf{A}$. The *i*-th row vector of $\mathbf{V}$ corresponds to d_i, and the 9-th row vector is the least squares solution of the homography parameter $\mathbf{h}$. Finally, the parameters in $\mathbf{H}$ are normalized so that $h_9 = 1$.

2.3.2. Estimation of Homography Matrix Using RANSAC

The RANSAC algorithm can estimate reasonable fitting parameters, even from data including outliers. The algorithm is performed as follows.

1. Select four optical flows randomly.

2. Calculate the homography matrix **H** by applying the DLT algorithm to the four optical flows.
3. Count the number of optical flows with back projection errors less than a certain value ε as follows:

$$\left(x' - \frac{h_1 x + h_2 y + h_3}{h_7 x + h_8 y + h_9}\right)^2 + \left(y' - \frac{h_4 x + h_5 y + h_6}{h_7 x + h_8 y + h_9}\right)^2 < \varepsilon. \tag{14}$$

The optical flows which satisfy Equation (14) are determined to be inliers, and the others are determined to be outliers.

4. Iterate the above steps from 1 to 3 for a certain time.
5. Determine the pre-optimal homography matrix that produces the most inliers.
6. Calculate the optimal homography matrix from the inliers of the pre-optimal homography matrix.

Most inliers originate from the camera motion, whereas most outliers originate from moving objects or false optical flows.

2.4. Extraction of Optical Flows on Moving Steps

The frame at $t = t$ is transformed on the basis of the optimal homography matrix. The image subtraction is performed between the transformed frame and the frame at $t = t + 1$. The moving steps appear as regions with high intensities in the subtraction image. These regions are extracted by a binarization operation followed by the closing and opening operations of the mathematical morphology for shape smoothing. The extracted regions are used as masks to select the optical flows on moving steps. A rectangular region of interest (ROI) of $H_{ROI} \times W_{ROI}$ pixels is set on the middle area of the frames to exclude unnecessary optical flows caused by non-interest objects such as people, as shown in Figure 3. The optical flows on the masks in the ROI are extracted to recognize the escalator.
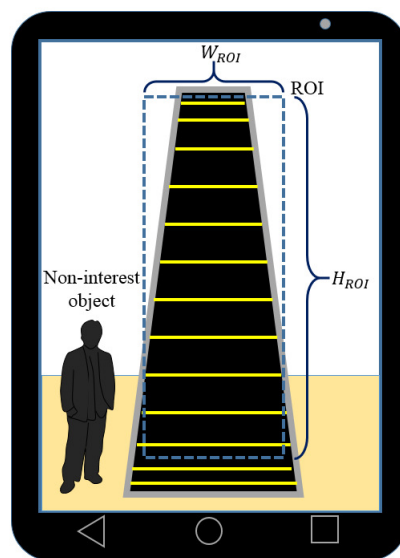

Figure 3. Region of interest (ROI) for escalator recognition.

2.5. Recognition of an Escalator

Escalators are categorized into the following four classes according to their movement directions:

1. Escalators going to upper floors (denoted by E_{TU})
2. Escalators going to lower floors (E_{TL})
3. Escalators coming from upper floors (E_{FU})
4. Escalators coming from lower floors (E_{FL})

The escalator classes are determined from the inliers optical flows. First, the camera motion vector is obtained by averaging the inliers optical flows. All the optical flows are subtracted by the camera motion vector. From the subtracted flows, the false optical flows with lengths more than L_{of}^u or less than L_{of}^l are eliminated. The final optical flows represent the movement direction of the steps of the escalator. If the movement direction is upward in the frame, the escalator is determined to be E_{TU} or E_{TL}. Otherwise, it is determined to be E_{FU} or E_{FL}. Note that E_{TL} and E_{FL} escalators produce upward and downward optical flows in video frames, respectively. Figure 4 shows the final optical flows on E_{TU} and E_{TL} escalators. White circles and red lines represent the corners and the optical flows, respectively. The E_{TL} and E_{TU} escalators produce upward optical flows.
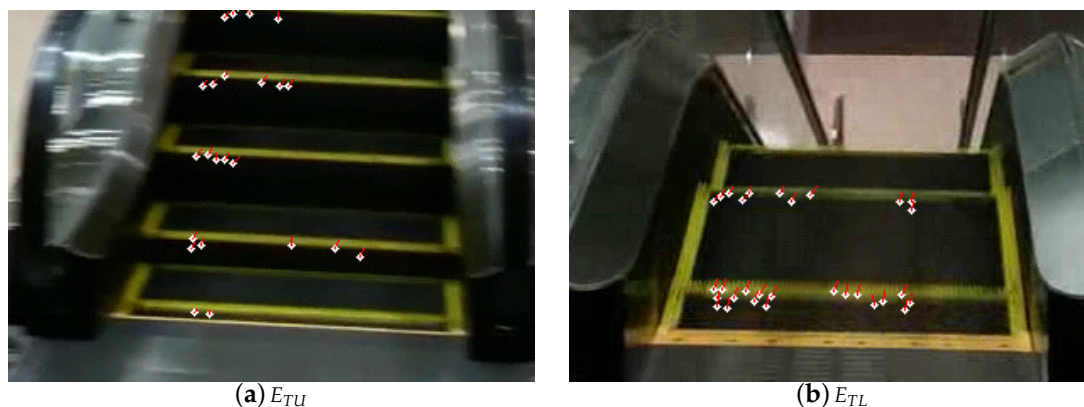

(**a**) E_{TU} (**b**) E_{TL}

Figure 4. The final optical flows on E_{TU} and E_{TL} escalators.

Next, further classification is performed on the basis of the numbers of steps observed in the frames. The step numbers are obtained from the numbers of groups where the vertical distance between two corner points is closer than D_{cp}. If the step numbers are larger than N_{step}, the escalator is determined to be E_{TU} or E_{FU}. Otherwise, it is determined to be E_{TL} or E_{FL}.

2.6. Notification to a User

Visually impaired users can use E_{TU} and E_{TL} escalators to move to upper and lower floors, respectively, whereas they cannot use E_{FU} and E_{FL} escalators. The system provides navigation sounds with higher and middle frequencies for E_{TU} and E_{TL}, respectively. The system also provides warning sounds with a lower frequency for E_{FU} and E_{FL}. The users can select navigation and warning sounds from several sound patterns and also adjust the sound frequencies beforehand so that they can distinguish the sounds effectively.

3. Experiments

3.1. Conditions

We used the Android smartphone Nexus 5 [51] with a Full High Definition touchscreen. In the corner detection process, the size of S_p was set to 3×3. Q_λ, d_{cd}, and N_{fn} were set to 0.05%, 5 pixels, and 5000, respectively. In the optical flow computation process, M_{rs} was set to 3 pixels. The ROI sizes H_{ROI} and W_{ROI} were set to 1024 and 120, respectively. In the escalator recognition process, L_{of}^u, L_{of}^l, D_{cp}, and N_{steps} were set to 8, 2, 24, and 6, respectively.

The system was evaluated using pre-recorded video frames taken at six points near six escalators as depicted in Figure 5. The smartphone was set at 3 or 5 m from the gates of the escalators and panned horizontally to include the whole of the escalator in the video frames. Figures 6 and 7 are sample escalators. We also used video frames from 24 scenes that did not include any escalators but included several moving objects such as humans, cars, and bikes.
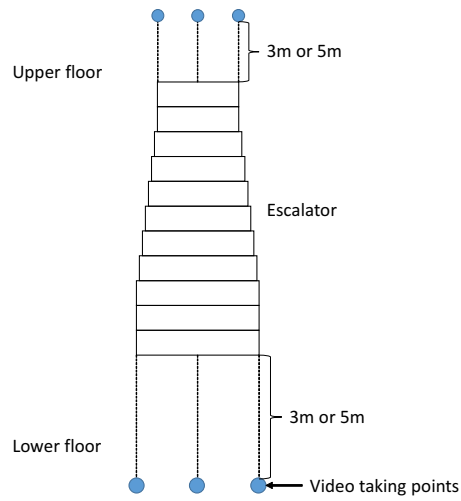
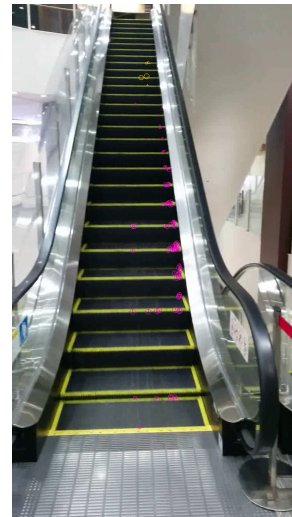**Figure 5.** Video taking points near an escalator.

Figure 6. E_{TU} (**left**) and E_{FU} (**right**) escalators observed at 3 m.
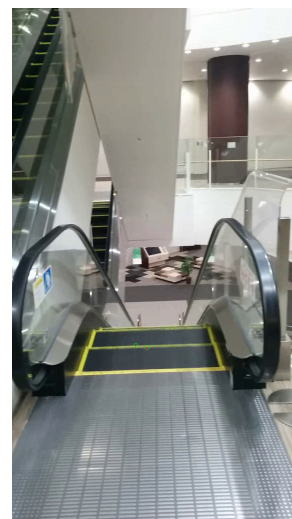

Figure 7. E_{FL} (**left**) and E_{TL} (**right**) escalators observed at 3 m.

3.2. Results

Tables 1 and 2 list the recognition results of escalators at 3 and 5 m, respectively. The system recognized 97% of E_{TU} and E_{FU} escalators correctly. In contrast, it failed to recognize 18% of E_{TL} and E_{FL} escalators. The E_{TL} and E_{FL} escalators were observed from upper floors, as shown in Figure 7. The system was not able to obtain a sufficient number of corner points and optical flows, which made recognition unstable. On the other hand, the system was able to correctly recognize all the videos of scenes that did not include escalators.

Table 1. Recognition results at 3 m.

		Output				
		E_{TU}	E_{TL}	E_{FU}	E_{FL}	Others
	E_{TU}	18	0	0	0	0
	E_{TL}	0	14	0	0	4
Input	E_{FU}	0	0	17	1	0
	E_{FL}	0	0	0	13	5
	Others	0	0	0	0	12

Table 2. Recognition results at 5 m.

		Output				
		E_{TU}	E_{TL}	E_{FU}	E_{FL}	Others
	E_{TU}	17	1	0	0	0
	E_{TL}	0	16	0	0	2
Input	E_{FU}	0	0	18	0	0
	E_{FL}	0	0	0	16	2
	Others	0	0	0	0	12

4. Discussion

The technical contribution of this paper is to adequately combine the image processing algorithms such as the corner detection, optical flow computation, and image registration. Although they are existing algorithms, the integrated method can do the task (i.e., escalator recognition), which has not been achieved by the previously proposed methods. In Section 3, we designed the experiments considering the variation of the relative distances and directions of users against escalators. The analysis results demonstrated that the proposed method was effective for escalator recognition.

The contribution from a welfare point of view is to be able to help visually impaired individuals use escalators that are representative dynamic objects in daily life. In this paper, we adopted the optical flow analysis to recognize escalators. This analysis method can be applied to other dynamic objects, and would make the lives of the individuals more convenient.

Our preliminary investigation revealed that many visually impaired individuals identified escalators by listening to mechanical sounds, then determined their movement directions by touching the belts. This can be dangerous. By using the proposed system, escalators can be recognized more safely.

The proposed system would not be able to work well in crowded environments such as stations in rush hour times, because other passengers on escalators would produce optical flows that are different from those of the steps. The different optical flows cause misrecognition. The current system cannot deal with this problem; therefore, users should determine the environments by hearing and select whether or not to use the system. In the future, this problem should be solved by eliminating optical flows in human regions extracted by human detection methods such as the histograms of oriented

gradients (HOG) technique [52]. In addition, some passengers may not like to have their pictures taken with smartphone cameras. It is not a technical issue, but social understanding is needed.

The proposed method assumes that there are salient features on the steps of escalators. Such features were detected by the corner detection method, and used for optical flow computation. In many countries, escalator steps would have such features to give cautions to passengers, but there are different escalators in the world. It is necessary to improve the system to recognize such escalators correctly.

The proposed method was implemented on a Nexus 5 smartphone, and the processing rate was approximately one frame per second. The processing speed should be increased to make the system more practical.

One of the authors is blind. The author mentioned that it was easy to take the pictures of escalators because the author can know the approximate positions by hearing the motor sounds. The author also mentioned that it was more important to know the moving directions of the steps without touching the belts. The author appreciated the proposed system to be able to compensate the hearing sense. These comments indicate that the system is effective in assisting the visually impaired.

In this paper, we performed only the technical experiments in Section 3. Although the blind author appreciated the effectiveness of the proposed system, it is not guaranteed that the proposed system is effective for all visually impaired individuals. However, the experimental results can imply that the system would be able to help many users find available escalators effectively. In this paper, we mainly proposed the system from the viewpoint of system development. In the future, we should evaluate its effectiveness with actual visually impaired individuals, especially blind people.

5. Conclusions

This paper proposed a smartphone-based assistive system to enable visually impaired individuals to use escalators safely. The system can detect escalators and determine their movement directions from optical flows in videos obtained by a smartphone camera. The system was evaluated in actual scenes that involved escalators and other objects. The experimental results demonstrate that the system is promising in term of helping visually impaired individuals use escalators.

Author Contributions: D.N. developed the algorithm, wrote the programs, acquired the data, conducted experiments, and analyzed the results. H.T. initiated and led the study and was the corresponding author of the manuscript. 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. WHO. 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. Leo, M.; Medioni, G.G.; Trivedi, M.M.; Kanade, T.; Farinella, G.M. Computer Vision for Assistive Technologies. *Comput. Vis. Image Underst.* **2017**, *154*, 1–15.
3. Bolgiano, D.; Meeks, E. A Laser Cane for the Blind. *IEEE J. Quantum Electr.* **1967**, *3*, 268.
4. 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, 31 January–2 February 1973; Volume 12, pp. 53–57.
5. Benjamin, J.M., Jr. The Laser Cane. *J. Rehabil. Res. Dev.* **1974**, *BPR 10-22*, 443–450.
6. Okayasu, M. Newly developed walking apparatus for identification of obstructions by visually impaired people. *J. Mech. Sci. Technol.* **2010**, *24*, 1261–1264.
7. Akitaseiko. 1976. Available online: http://www.akitaseiko.jp (accessed on 6 January 2017).

8.　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.

9.　Dang, Q.K.; Chee, Y.; Pham, D.D.; Suh, Y.S. A Virtual Blind Cane Using a Line Laser-Based Vision System and an Inertial Measurement Unit. *Sensors* **2016**, *16*, 95.

10.　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.

11.　Ju, J.S.; Ko, E.; Kim, E.Y. EYECane: Navigating with camera embedded white cane for visually impaired person. In Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility, Pittsburgh, PA, USA, 25–28 October 2009; pp. 237–238.

12.　Vera, P.; Zenteno, D.; Salas, J. A smartphone-based virtual white cane. *Pattern Anal. Appl.* **2014**, *17*, 623–632.

13.　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.

14.　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.

15.　Huang, H.C.; Hsieh, C.T.; Yeh, C.H. An Indoor Obstacle Detection System Using Depth Information and Region Growth. *Sensors* **2015**, *15*, 27116–27141.

16.　McDaniel, T.; Krishna, S.; Balasubramanian, V.; Colbry, D.; Panchanathan, S. Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. In Proceedings of the International Workshop on Haptic Audio Visual Environments and Games, Ottawa, ON, Canada, 18–19 October 2008; pp. 13–18.

17.　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; Volume IV, pp. 584–587.

18.　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.

19.　Balakrishnan, G.; Sainarayanan, G.; Nagarajan, R.; Yaacob, S. Wearable Real-Time Stereo Vision for the Visually Impaired. *Eng. Lett.* **2007**, *14*, 1–9.

20.　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, AZ, USA, 7–10 November 2010; pp. 1173–1178.

21.　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, Berkeley, CA, USA, 12 July 2011; pp. 1–6.

22.　Helal, A.S.; Moore, S.E.; Ramachandran, B. Drishti: An integrated navigation system for visually impaired and disabled. In Proceedings of the Fifth International Symposium on Wearable Computers, Zurich, Switzerland, 7–9 October 2001; pp. 149–156.

23.　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.

24.　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.

25.　Dunai, L.D.; Pérez, M.C.; Peris-Fajarnés, G.; Lengua, I.L. Euro Banknote Recognition System for Blind People. *Sensors* **2017**, *17*, 184.

26.　Ulrich, I.; Borenstein, J. The GuideCane-Applying Mobile Robot Technologies to Assist the Visually Impaired. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2001**, *31*, 131–136.

27.　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.

28.　Cloix, S.; Bologna, G.; Weiss, V.; Pun, T.; Hasler, D. Low-power depth-based descending stair detection for smart assistive devices. *EURASIP J. Image Video Proc.* **2016**, *2016*, 33.

29. Saegusa, S.; Yasuda, Y.; Uratani, Y.; Tanaka, E.; Makino, T.; Chang, J.Y. Development of a Guide-Dog Robot: Leading and Recognizing a Visually-Handicapped Person using a LRF. *J. Adv. Mech. Des. Syst. Manuf.* **2010**, *4*, 194–205.

30. Peng, E.; Peursum, P.; Li, L.; Venkatesh, S. A smartphone-based obstacle sensor for the visually impaired. In Proceedings of the Ubiquitous Intelligence and Computing, Xi'an, China, 26–29 October 2010; pp. 590–604.

31. Dumitraş, T.; Lee, M.; Quinones, P.; Smailagic, A.; Siewiorek, D.; Narasimhan, P. Eye of the Beholder: Phone-based text-recognition for the visually-impaired. In Proceedings of the 10th IEEE International Symposium on Wearable Computers, Montreux, Switzerland, 11–14 October 2006; pp. 145–146.

32. Tekin, E.; Coughlan, J.M.; Shen, H. Real-time detection and reading of LED/LCD displays for visually impaired persons. In Proceedings of the 2011 IEEE Workshop on Applications of Computer Vision (WACV), Kona, HI, USA, 5–7 January 2011; pp. 491–496.

33. Zhang, S.; Yoshino, K. A braille recognition system by the mobile phone with embedded camera. In Proceedings of the Second International Conference on Innovative Computing, Information and Control, Kumamoto, Japan, 5–7 September 2007; pp. 223.

34. American Foundation for the Blind. Estimated number of adult Braille readers in the United States. *J. Vis. Impair. Blind.* **1996**, *90*, 287.

35. Al-Doweesh, S.A.; Al-Hamed, F.A.; Al-Khalifa, H.S. What Color? A Real-time Color Identification Mobile Application for Visually Impaired People. In Proceedings of the HCI International 2014-Posters Extended Abstracts, Heraklion, Crete, 22–27 June 2014; pp. 203–208.

36. Matusiak, K.; Skulimowski, P.; Strumillo, P. Object recognition in a mobile phone application for visually impaired users. In Proceedings of the 6th International Conference on Human System Interaction (HSI), Gdansk, Poland, 6–8 June 2013; pp. 479–484.

37. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.

38. Rosten, E.; Drummond, T. Machine learning for high-speed corner detection. In Proceedings of the Computer Vision–ECCV 2006, Graz, Austria, 7–13 May 2006; pp. 430–443.

39. Ivanchenko, V.; Coughlan, J.; Shen, H. Crosswatch: A camera phone system for orienting visually impaired pedestrians at traffic intersections. In Proceedings of the International Conference on Computers for Handicapped Persons, Linz, Austria, 9–11 July 2008; pp. 1122–1128.

40. Ivanchenko, V.; Coughlan, J.; Shen, H. Real-time walk light detection with a mobile phone. In Proceedings of the Computers Helping People with Special Needs, Vienna, Austria, 14–16 July 2010; pp. 229–234.

41. 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.

42. Mocanu, B.; Tapu, R.; Zaharia, T. An Obstacle Categorization System for Visually Impaired People. In Proceedings of the 2015 11th International Conference on Signal-Image Technology Internet-Based Systems (SITIS), Bangkok, Thailand, 23–27 November 2015; pp. 147–154.

43. Mocanu, B.; Tapu, R.; Zaharia, T. When Ultrasonic Sensors and Computer Vision Join Forces for Efficient Obstacle Detection and Recognition. *Sensors* **2016**, *16*, 1807.

44. 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.

45. Gonzalez R.C.; Woods, R.E. *Digital Image Processing*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2008; pp. 657–661.

46. Shi, J.; Tomasi, C. Good features to track. In Proceedings of the 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 21–23 June 1994; pp. 593–600.

47. Lucas, B.D.; Kanade, T. An iterative image registration technique with an application to stereo vision. *IJCAI* **1981**, *81*, 674–679.

48. Bouguet, J.Y. Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm. *Intel Corp.* **2001**, *5*, 4.

49. Solem, J.E. *Programming Computer Vision with Python-Tools and Algorithm for Analyzing Images*; O'Reilly Media, Inc.: Newton, MA, USA, 2012; pp. 55–58.

50. Szeliski, R. *Computer Vision-Algorithms and Applications*; Springer: London, UK, 2010.

51. Google. Nexus 5. Available online: https://www.google.com/nexus/5x/ (accessed on 1 May 2016).

52. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20–26 June 2005; Volume 1, 886–893.