Camassia: Monocular Interactive Mobile Way Sonification

Sebastian Ritterbusch[1,2](✉) and Gerhard Jaworek[3]

[1] iXpoint Informationssysteme GmbH, Ettlingen, Germany
sebastian.ritterbusch@ixpoint.de
[2] VWA-Hochschule, Stuttgart, Germany
[3] Study Centre for the Visually Impaired, Karlsruhe Institute of Technology, Karlsruhe, Germany
gerhard.jaworek@kit.edu

Abstract. Real-time camera image analysis informs the walking person about the way in front by sonification. For visually impaired people, this opens a new way to experience their surroundings and move more safely outdoors. We extended an image analysis workflow from autonomous robots to human interaction, adaption to alternating way appearance, and various sonification options. The solution is available on off-the-shelf smartphones.

Keywords: Sonification · Navigation · Assistive system · Blind
Free way detection

1 Introduction

Autonomous and self-determined mobility is essential to our life. Visual impairments are a disadvantage with respect to recognition of the way ahead. Systems to detect obstacles like the white cane have been proven to be very effective and helpful. Many technical solutions use various sensors to extend the detected area, or to solve related problems to mobility. This often leads to the use of specialized hardware that can be quite costly, or solutions that do not leave experimental status, and are therefore not generally available.

The presented approach uses the camera and inertial sensors of an off-the-shelf smartphone to enhance the perception of the way in front, by sonification. By a low-latency approach from camera analysis to auditory feedback, the solution gives immediate feedback to the user rotating the camera in order to scan the surroundings.

The adaption of the visual recognition algorithm of a competition winning autonomous robot relies on the user to decide which way to go, instead of let the person being controlled by directions of an expert system.

K. Miesenberger and G. Kouroupetroglou (Eds.): ICCHP 2018, LNCS 10897, pp. 12–18, 2018.
https://doi.org/10.1007/978-3-319-94274-2_2

2 Problem Statement and Related Work

There are several approaches to aid the navigation of visually impaired humans based on smartphones, but they either require detectable markers in the environment [1], or a priori an accurate map [2]. In [3], one of the authors applied sonification methods in combination with a force-feedback device to locate objects in virtual space.

In contrast to many obstacle detecting solutions, the presented approach follows the concept of [4] to identify free space instead. The idea in the vision algorithm [5] of the KAMARO robot car which won the European ROBOTOUR 2015 competition is fundamental to this work. It seems relatively simple compared to the work presented in [6], but instead of detecting the whole road and vanishing point, it aims at robustly and very efficiently computing a one-dimensional angle distribution of free passage probabilities, for interactive real-time and interactive use. A one dimensional density function is very useful for interactive sonification for exploration as in [7], as found with participation of one of the authors. It has to be noted, that the use of the audio channel for wayfinding is generally disadvantageous from the point of cognitive load [8] for the blind, but the choice of the smartphone platform does not offer other appropriate fast updating communication channels, yet.

The aim of the solution is to provide live acoustic feedback of the camera image to aid visually impaired persons in finding passable ways.

3 Methodology

The basic concept is based on the observation of Fürst in [5,9] that pavements generally have low color saturation compared to the surroundings. In contrast to a fixed camera setup on a robot car, first of all, the image of a human controlled camera has to be rectified. Additionally, the method was extended to adapt to general and changing soil appearances.

Due to the perspective mapping of the surroundings onto the camera image, the image analysis yields a one-dimensional angle distribution of free passage probabilities. Figure 1 illustrates the steps of the image analysis workflow: In the first step, an image is taken from the camera. Since we cannot expect the blind user to hold the camera straight, the image then is rectified using the Inertial Measurement Unit (IMU) of the smartphone. Also, image data above the horizon is discarded. In the next step, the remaining image is color adjusted with respect to the last color calibration by the user. The result is then analyzed for way components and finally summarized in the free way probability distribution, which will be shown as a bar on top of the screen.

This result is mapped onto a two octave spanning stereophonic real-time audio synthesizer, giving immediate feedback to the perceived floor image. The distribution is mapped onto 24 half-tones with its intensity representing the assumed probability. Alternative sonification modes allow to replace the synthesizer by stereophonic white noise, or to reduce the sound to local maxima in the density function to reduce the sound, and therefore the cognitive load, even more.

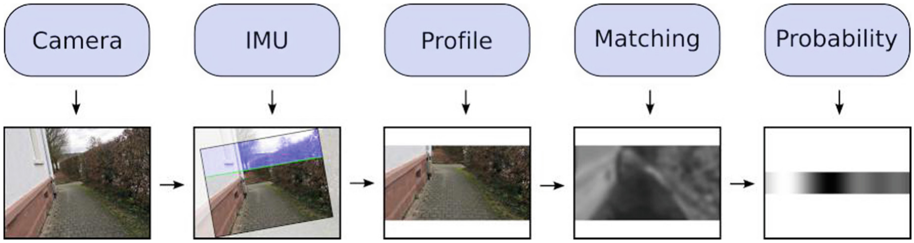

Fig. 1. From the camera the image is rectified and cut, the image is transformed following the color profile, analyzed for its way components and finally transformed to the free way probability distribution. (Color figure online)

The standard sonification method is redundant by purpose, both using sound frequency and panorama to reflect the probability density. This way, the system remains usable with monophonic smartphone speakers. On the other hand, it is offering additional panoramic information when a stereophonic headset is used.

Due to the challenging representation of visual components via polyphonic acoustic representations, the solution is supplemented by an introductory narrated tutorial and a more detailed tutorial introducing and explaining the sound options.

4 Results

While the original algorithm was aimed at pavements, its extension to general soil appearance turned out to be quite robust, see Fig. 2 for two examples. It is quite successful in detecting earth paths as well as colored pavements, as well as colored indoor carpeting. The solution reaches its limits in the case of sidewalks if they are paved in the same way as the street, and is of very limited use when crossing a street.

While the solution should be robust to varying illumination of the path, this has limits due to non-linearity of the camera, and may also lead to false detection of pure white or black areas. Further refinement of the algorithm will take this into account.

On current devices, the algorithm easily analyzes 30 images in a second, leading to a very fast response in the audio output. Due to device based audio buffering and necessity of acoustic fading of sounds, there is short delay from detection to acoustic response lower than a tenth of a second.

Besides the contributions and first hand experience of one of the authors with respect to blindness, twelve candidates being blind or mobility trainers were invited to participate in a test before publication of the software, of which nine persons then participated in the test. They were given an questionnaire covering the information about the test person, usability, application fields, the human computer interface, and the overall impression. Five of the testers gave answers to at least parts of the questionnaire.

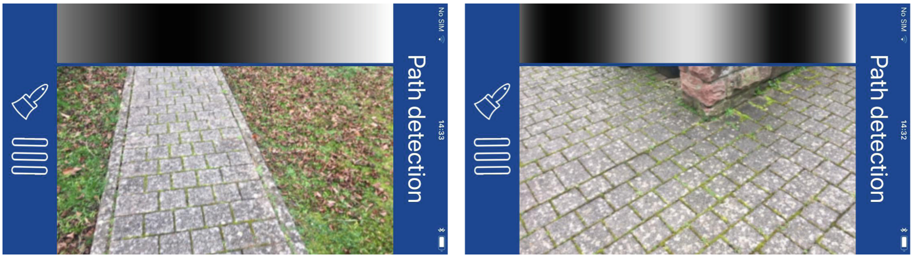

Fig. 2. Two screenshots of live path detection show the image of the camera in the lower part, a grayscale bar in the upper part represents the probability density function of assumed free space represented in darker shades. The first screenshot shows how a straight path is detected. The second screenshot shows a branching which two branches were accurately detected. The brush icon as well as the pause or play button on headsets calibrates the color of the way, the three bar icon leads to the configuration. (Color figure online)

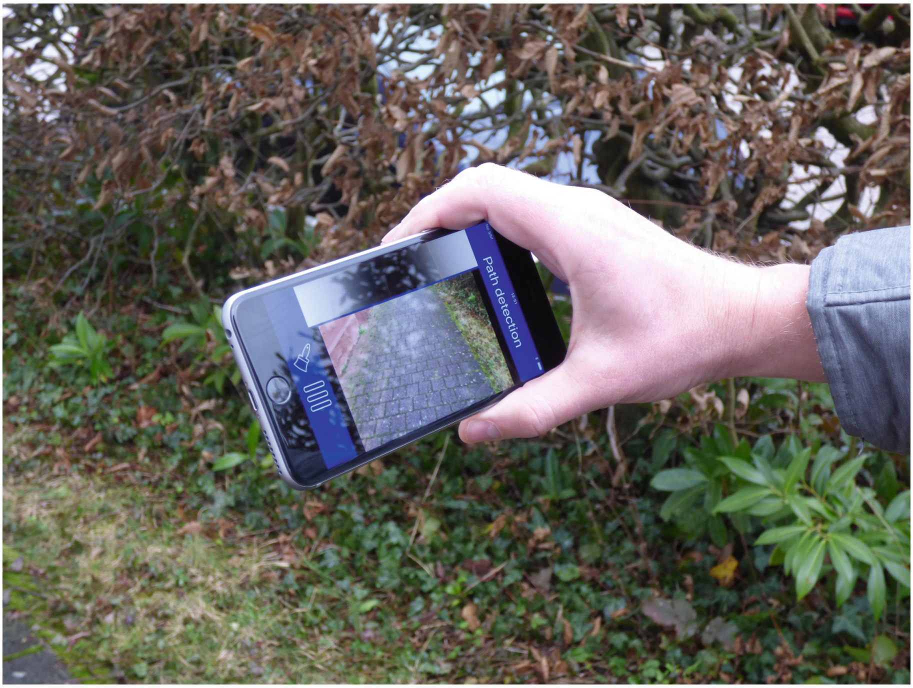

Fig. 3. For way detection, the smartphone should be held in front of the person, tilted downwards, slightly. Photo by H. Kucharek.

Fig. 4. The detection algorithm was found to be very effective for paths in parks, as well as pavements enclosed by grass. Photo by H. Kucharek.

As expected from [8], the choice of sonification for representing the free space probabilities understandingly was not enthusiastically received by visually impaired testers. Yet, the applicability for certain situations was both confirmed by blind users and mobility trainers. There was a strong demand for additional training material with sound examples, which was missing in the test version, and was added in the published version.

A useful example was found in identifying side ways that are not easily detectable by the white cane, as illustrated in Fig. 2. Also, the immediate acoustic response in advance to passing cyclists was very well received. The indoor use of the solution was an unexpected response. Another response was the need for more training, especially to take advantage of the immediate response to moving the camera. The use of bone-conduction earphones was preferred to the use of the internal smartphone speaker due to the panoramic response.

In practical use, the solution was efficient both in landscape and portrait mode. While the landscape mode is advised due to a larger field of view, a swinging motion of the camera to the left and right like the white cane, overcame the limitation. Best results where achieved for paved paths enclosed with vegetation and in parks, as illustrated in Figs. 3 and 4. As proposed in the supplied training, additional wide-angle lenses can strongly increase the field of view, but it turned out, that they are not easily attached to the smartphone by a blind person.

The use of a narrated tutorial was well received, stating that it would hardly possible to understand what the solution does without such a short introduction. Following the feedback from the testers, this tutorial was extended by giving more examples on where the App can be used.

5 Conclusion

The adaption and extension from an image analysis workflow from robotics in combination with various sonification modes led to an applicable solution for experienced blind users, as an addition to the concurrent use of the well-established white cane or guide dogs. To the knowledge of the authors, it is the first publicly obtainable solution dedicated to general and interactive path detection on the smartphone.

It is expected to be a starting point for real-time mobile camera-based path detection immediately communicated to the user, and will be followed by more sophisticated concepts based on Artificial Intelligence or more intensive pattern recognition due to the availability of computing power on the smartphone. Also, the future support of external haptic devices, such as vibration bands for arm, foot or waist as an alternative to sonification should improve the cognitive load.

The solution is available for general use on iTunes Store.

References

1. Coughlan, J., Manduchi, R.: A mobile phone wayfinding system for visually impaired users. Assist. Technol. Res. Ser. **25**(2009), 849 (2009)
2. Chen, D., Feng, W., Zhao, Q., Hu, M., Wang, T.: An infrastructure-free indoor navigation system for blind people. In: Su, C.-Y., Rakheja, S., Liu, H. (eds.) ICIRA 2012. LNCS (LNAI), vol. 7508, pp. 552–561. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33503-7_54
3. Ritterbusch, S., Constantinescu, A., Koch, V.: Hapto-acoustic scene representation. In: Miesenberger, K., Karshmer, A., Penaz, P., Zagler, W. (eds.) ICCHP 2012. LNCS, vol. 7383, pp. 644–650. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31534-3_94
4. Schauerte, B., Koester, D., Martinez, M., Stiefelhagen, R.: Way to go! Detecting open areas ahead of a walking person. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8927, pp. 349–360. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16199-0_25
5. Fürst, M.: Detecting drivable regions in monocular images. Robotour 2015: Autonomous Robot in Parks and Urban Regions (2015)
6. Kong, H., Audibert, J.Y., Ponce, J.: General road detection from a single image. IEEE Trans. Image Process. **19**(8), 2211–2220 (2010)
7. Wörtwein, T., Schauerte, B., Müller, K., Stiefelhagen, R.: Mobile interactive image sonification for the blind. In: Miesenberger, K., Bühler, C., Penaz, P. (eds.) ICCHP 2016. LNCS, vol. 9758, pp. 212–219. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41264-1_28

8. Martinez, M., Constantinescu, A., Schauerte, B., Koester, D., Stiefelhagen, R.: Cognitive evaluation of haptic and audio feedback in short range navigation tasks. In: Miesenberger, K., Fels, D., Archambault, D., Peňáz, P., Zagler, W. (eds.) ICCHP 2014. LNCS, vol. 8548, pp. 128–135. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08599-9_20
9. Fürst, M.: Probabilistische Robotik, conversation with S. Ritterbusch in the Modellansatz Podcast, Episode 95, Department of Mathematics, Karlsruhe Institute of Technology (KIT) (2016). http://modellansatz.de/probabilistische-robotik