Independent way-finding for visually impaired users through multi-sensorial data analysis on mobile devices

Ahmetovic Dragan - dragan.ahmetovic@unimi.it
Dipartimento di Informatica e Comunicazione, Università degli Studi di Milano, Italy

I. Introduction

For users with sensory or mobility impairments the evolution of mobile devices introduces new accessibility issues. For instance, touch screens, lacking tactile cues, are hard to use by blind users without additional feedback. These problems are being mitigated through new interaction paradigms, such as screen readers: software that describe the currently selected elements on the interface through vocal cues.
Modern mobile devices, with suitable accessibility features, can be used to overcome limits caused by disabilities and contribute towards a greater independence of disabled users. A dangerous task for an unassisted blind user, such as orientation in unfamiliar environment, can be tackled through the usage of data derived from sensors available on recent mobile platforms (video cameras, GPS, accelerometers).

II. Problem statement

Accessing visual information (colors, lights and shapes) is one of the greatest issues for blind users. This problem arises when interacting with lightsources, objects lacking distinguishing tactile features (boxes, pictures) or unreachable objects (landmarks, signs). While proprietary hardware solutions exist for some issues (e.g.:color/light detection), the same functionalities can also be developed for less bulky, multi-purpose mobile platforms.
The recognition of complex graphical elements is possible through computer vision techniques, executed on either mobile device or remote platforms. Remote detection benefits from higher computational power, however, due to the data transfer times, remote recognition can result in detection delays, unsuitable for time-critical applications. In these cases onboard video stream-based detection is preferred.
Independent way-finding is another major problem for blind users, even more in unexplored surroundings. Unassisted navigation is hazardous due to obstacles and vehicles. While GPS can be used for coarse navigation, precise orientation needs a different approach. Computer vision-assisted way-finding (using accelerometer and other sensors' data through Sensor Fusion techniques) has already been considered [2], and usage of information from surrounding pervasive devices and internet is a possible extension of this work. While traffic lights [3] and pedestrian crossing [2] detection is possible, the recognition of other navigation cues, such as road and shop signs, house numbers and other textual data, is another considered step towards a safer and independent mobility for the visually impaired users.
The development of custom interfaces, designed to remove the interaction obstacles while preserving the user's safety, is another important issue. During mobility, the visually impaired user's hearing sense is crucial for danger avoidance. With one user's hand occupied maneuvering a white cane or holding a guide dog's leash, one handed interaction with limited or no audio output (e.g.: haptic interaction techniques) have to be considered.

III. Related work

Accesibility issues for visually impaired users have arised due to the shift of the interaction paradigm from a tactile approach towards a graphically oriented one (keyboard to touchscreen). Screen readers, such as Voiceover on iOS, address this problem. Exploration through Screen readers, however, can be time consuming and the content may also be missed if too small. Ahmed et al. [1] address the issue, limited to the web browsing on touch interfaces. The proposed solution is effective for tabular data and suggests the usage of directional swiping gesture-based navigation.
Alternative text input methods, in contrast with typical software qwerty keyboards which are hard to use by visually impaired subjects, are proposed in [6]. The solution, named "TypeInBraille", relies for the text entry on sequential gestures which imitate the rows of Braille characters.
Unobtrusive interaction paradigms during the navigation are preliminarily explored in [2]. Short audio cues inform the user about the position of detected zebra crossings, minimizing the loss of attention. Test users have shown interest towards further research on safety-aware user interfaces.
In the field of assistive technologies, indoors navigation through Wi-Fi and RFID beacons is suggested by Li et al. [5]. This approach, however, requires the usage of specialized facilities in the environment of interest.
Angin et al. [3] consider the traffic lights recognition through cloud computing. Response times are lower than a second under the premise that Wi-Fi connection is available.
Ivanchenko et al. [4] deal with computer vision-based detection of zebra crossings on mobile devices. Since the position of the zebra crossing is not computed, the alignment of the user towards the crosswalk is not measured and improved with feedback. This issue is addressed in [2]: through the

978-1-4244-9529-0/12/$31.00 ©2012 IEEE

usage of accelerometers, the zebra crossing's position with respect to the user is determined and the user is assisted during the crossing with unobtrusive vocal messages.

IV. Expected and Achieved Contributions

The overall expected contribution is a comprehensive navigation and orientation platform for visually impaired users on mobile devices. The solution is expected to provide a coarse long range GPS navigation while computer vision and sensor-driven orientation will be adopted for close range guidance. As shown in [2], the zebra crossing detection has already been achieved (See fig. 1) and detection of other cues such as traffic and shop signs, house numbers and textual information is being considered. Additional integration with network and surrounding pervasive data sources will provide information sharing between users as well as data regarding surrounding points of interest and available location based services. Non-invasive and consistent interaction paradigms based on haptic and audio interaction will assist the user during the navigation, keeping in mind the safety concerns that can arise during the navigation in unexplored environments.

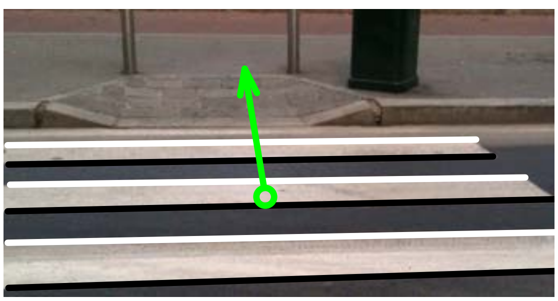

Figure 1. Detected zebra crossing with position and direction

V. Approach

The development of effective user interfaces depends on the definition of appropriate interaction paradigms. Interaction paradigms are consistent sets of input/output methodologies, defined according to the user's capabilities and conveyed data. Context appropriateness (the suitability of interaction paradigms in a usage context) is also important for the definition of interaction paradigms. Haptic interaction techniques in noisy environments, as suggested in section II are an example of context appropriate interfaces.

For Computer vision-based object recognition, analytical and machine learning approaches will be considered. The analytical approach relies on the analysis of a manually selected set of cues. This approach is more appropriate for the detection of objects having a low number of distinctive features [2]. The machine learning approach consists in an initial automated training during which suitable features are selected from a pre-classificated set of images. For objects with high number of less defining features (e.g: face detection [7]), this approach is preferred. SVM, Haar features and AdaBoost boosting techniques [7] will be considered.

In [2], the usage of data from different sensors (besides video camera) is shown to impact positively both the quality and the performance of the recognition. Through the usage of Sensor Fusion techniques, contextual information from other data sources can be used for an accurate detection and a faster prediction of the position of the detected object. For example, after an object has been correctly detected, gyroscope data can be used to predict the position of the object after the user's next movement, without further executions of the computationally expensive recognition stage.

The results are analyzed through user-driven and computer-driven evaluations. Through quantitative and qualitative measurements of the interaction between users and the system, the user-driven evaluation measures the effectiveness of the solution, the appropriateness of the interface and the fatigue caused by the interaction. Computer-driven evaluation assesses the quality of computer vision and sensor fusion-based object recognition through computation time, precision (false positives ratio) and recall (false negatives ratio) metrics. Due to the safety concerns, the precision (false positives ratio) and computation time are the more delicate metrics. Generally, the results are required to yield a perfect precision score, with near real-time time detection.

Acknowledgments

I would like to thank my advisor dr. Sergio Mascetti, professor Claudio Bettini and dr. Cristian Bernareggi for their assistance and insight.

References

[1] Ahmed, F., Islam, M.A., Borodin, Y., and Ramakrishnan, I.V. Assistive web browsing with touch interfaces. In *Proc. of Int. Conf. on Computers and accessibility (Sigaccess)*. ACM, 2008.

[2] Ahmetovic, D., Bernareggi, C., and Mascetti, S. Zebralocalizer: identification and localization of pedestrian crossings. In *Proc. of 13th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services*. ACM, 2011.

[3] Angin, P., Bhargava, B., and Helal, S. A mobile-cloud collaborative traffic lights detector for blind navigation. In *Proc. of Int. Conf. on Mobile Data Management*. IEEE, 2010.

[4] Ivanchenko, V., Coughlan, J., and Shen, H. Staying in the crosswalk: A system for guiding visually impaired pedestrians at traffic intersections. In *Assist technol Res Ser.* IOS, 2009.

[5] Li, B., Ramsey-Stewart, E., Johar, K., Woo, D., and Rizos, C. More freedom for the blind and vision impaired-a proposed navigation and information system. In *IGNSS Symposium*. ACM, 2009.

[6] Mascetti, S., Bernareggi, C., and Bellotti, M. Typeinbraille: A braille-based typing application for touchscreen devices. Demo in Int. Conf. on Computers and accessibility (Sigaccess), 2011.

[7] Viola, P. and Jones, M. Robust real-time face detection. In *Int.l Journal of Computer Vision (IJCV)*. Springer, 2004.