Non-visual-cueing-based Sensing and Understanding of Nearby Entities in Aided Navigation.

Juan Diego Gomez
University of Geneva
Computer Science Department
Geneva, Switzerland
+41 (22) 379 0152
juan.gomez@unige.ch

Guido Bologna
University of Geneva
Computer Science Department
Geneva, Switzerland
+41 (22) 379 0152
guido.bologna@unige.ch

Thierry Pun
University of Geneva
Computer Science Department
Geneva, Switzerland
+41 (22) 379 0152
thierry.pun@unige.ch

ABSTRACT

Exploring unfamiliar environments is a challenging task in which additionally, unsighted individuals frequently fail to gain perception of obstacles and make serendipitous discoveries. This is because the mental depiction of the context is drastically lessened due to the absence of visual information. It is still not clear in neuroscience, whether stimuli elicited by visual cueing can be replicated by other senses (cross-model transfer). In the practice, however, everyone recognizes a key, whether it is felt in a pocket or seen on a table. We present a context-aware aid system for the blind that merges three levels of assistance enhancing the intelligibility of the nearby entities: an exploration module to help gain awareness of the surrounding context, an alerting method for warning the user when a stumble is likely, and, finally, a recognition engine that retrieves natural targets previously learned. Practical experiences with our system show that in the absence of visual cueing, the audio and haptic trajectory playback coupled with computer-vision methods is a promising approach to depict dynamic information of the immediate environment.

Categories and Subject Descriptors

K.4.2 [**computers & society**]: Social Issues – *Assistive technologies for persons with disabilities, Handicapped persons/special needs.*

Keywords

Context-aware aid, visually impaired, visual cueing, assistance.

1. INTRODUCTION

According to the World Health Organization, 285 million people are visually impaired worldwide. In this context, the research community is increasingly focusing on the development of assistive technology. In fulfilling such purpose the study of strategies to depict visual information to the blind has attracted significant interest. Notwithstanding, the representation of dynamic environmental visual cues through assistive systems is lagging behind the growing requirements of the blind community for independent living. This work targets to efficiently convey visual information hardly attainable otherwise. We propose a model for assistive navigation/exploration (fig.1) aimed at fostering awareness of context by increasing the intelligibility of nearby entities and easing the interpretation of spatial relations:

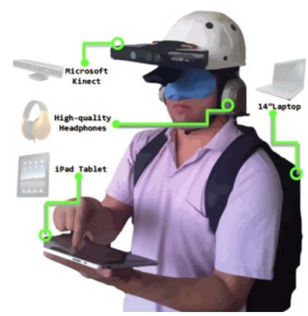

Figure 1. *The assistance system based on context-aware computing presented in this work.*

I. An exploration module that makes it possible for the users to touch with their fingers through an iPad, the whole 3D-image captured by a range camera in real time. The color and position of touched points are encoded in instruments sounds and sonic effects, respectively [2]. This module exploits the audio and haptic trajectory feedback to convey significant visual cues. Particularly, the use of spatialized sound allows understanding of spatial relations as sounds are perceived from 3D locations.

II. An alerting method based on range-imaging processing that prevents the user from stumbling by informing about unexpected entities lying on his way and potentially leading to a fall. At some point, this algorithm can also predict the trajectory of detected obstacles to keep/suspend a warning. This method allows the blind to find a clear path in the interest of safe navigation.

III. A recognition engine that uses state-of-the art object recognition methods to learn natural objects. There is a training phase supported by tracking and bootstrapping methods followed by an online searching process. This latter informs the user about the presence of learned objects in real time during exploration, if any. Unlike I and II, this engine is able to perceive and depict complex visual information in a much higher level of abstraction.

ASSETS'12, October 22–24, 2012, Boulder, Colorado, USA.
ACM 978-1-4503-1321-6/12/10.

2. EXPLORATION MODULE

Presenting nontextual or dynamic information through no heavily visual methods continues to be a challenge in aided navigation. In our system, pictures of a real-time 3D video stream are made possible to be touched with the fingers over an iPad. We convey the visual information of touched points by encoding its color and actual spatial position respectively, into instruments sounds and sonic effects (i.e. spatialization of sound) [3]. Ideally, our interface (fig. 2) turns the iPad into a door for the user to enter and sense the environment.

In general, when the user wants to explore the nearby space, (s)he will rapidly scan the touchpad with one or more fingers; the principle here is that finger movements will replace eye movements and more information from the image as a whole is made accessible in parallel. This provokes a proactive interaction to selectively explore, to discover points of interest, make comparisons, and, furthermore, enjoy a greater sense of independence.

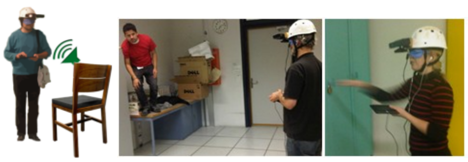

Figure 2. *Spatialization of sound is meant to create the illusion framed in this figure (leftmost). In our work, this allows to let the user hear the sound of a touched point in the picture (iPad) like it is coming from its counterpart point in the real world. In this example the user is touching (within the image displayed on the iPad he holds) the right-up corner of the chair's back. The others in this figure are random pictures of experimental sessions.*

3. ALERTING METHOD

The alerting method is range-imaging-based algorithm whose purpose is to warn the visually impaired user when a threatening situation is likely to happen as consequence of an unexpected obstacle. Once a warning has been launched, the user is expected to timely stop for a potential stumble be avoided. This also allows the blind to find a safe, clear path to advance through. Our guarding algorithm will run nonstop as long as the user is exploring. This user therefore, will be able to focus on gaining a context insight by means of the haptic-based interface, rather than minding his step.

To test this algorithm, we stuck the camera of our system on top of a rolling chair's back, which in turn carried the rest of the system on its seat. A person leaned on the chair arms so as to roll it down a corridor (with plenty of obstacles) while pushing with his walk. The end result is a video that shows the performance of our method: http://youtu.be/X426HAZaiYQ

4. RECOGNITION ENGINE

Based on the strategy proposed by Kalal et al [1], we have implemented a detecting-and-tracking hybrid method for learning the appearance of natural objects in unconstrained video streams. Thus, we allow visually impaired users gaining awareness of certain objects they otherwise could fail to perceive, or simply need others help to do so. Notice that sometimes looking for an object (e.g. a fallen object) may end up in an embarrassing situation that might lower their feeling of dignity. In general, we look forward to letting an unsighted individual be aware of serendipitously encounters such as a person on his way to the toilet but also, conscious searches such as for a telephone, an exit, a trash can etc. Our recognition engine permits: (1) Learn the object during a tracking phase. (2) When the learning meets an end, a sighted user provides the name of the object. (3) When the navigation/exploration task is being performed (fig. 3), the unsighted user is notified about the presence of the learned object through an earcon, audio icon or simply spelling out the name every time following detection.

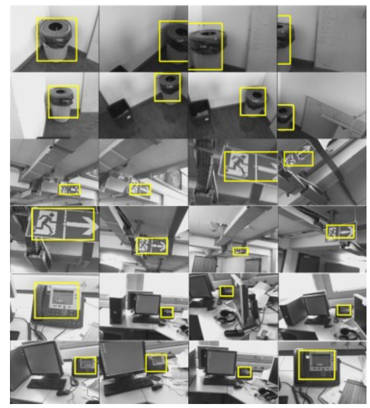

Figure 3. *Samples of real time detections of our recognition engine for three objects sequences. The yellow square represents the area were the object is thought to be, when detected. Rows 1 and 2: a trash can. Rows 3 and 4: an exit way. Rows 5 and 6: a telephone. These objects were selected as common targets by legally blind users in a survey.*

5. CONCLUSIONS

We presented a mobility aid system for unsighted individuals to allow them sensing and understanding a great deal of visual information about their immediate environment. The three levels of assistance that our context-aware system provides were described in this paper: an exploration module, an alerting method and a recognition engine. This system assesses the actual feasibility of assistive technologies to efficiently convey compacted visual information hardly attainable otherwise. We have showed that the use of the audio and haptic trajectory feedback is an effective method to encode visual features such as color and depth which in turn, make the environment more intelligible.

6. REFERENCES

[1] Kalal, Z., Matas, J. and Mikolajczyk, K (2010). P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints. In proceedings of CVPR'10, International Conference on Computer Vision and Pattern Recognition. San Francisco CA, US.

[2] Bologna, G., Deville, B., Gomez, J. and Pun, T (2010). Toward local and global perception modules for vision substitution. Neurocomputing, Vol. 74, Issue. 8, Pag. 1182–1190, Pub. date: October 2010.

[3] Ordonez, C., Navarun, G. and Barreto, A (2002). Sound spatialization as a navigational aid in virtual environments. In proceedings of CSI, The 6th Computer Security Institute Conference. Orlando Fl, US.