2010 12th International Workshop on Cellular Nanoscale
Networks and their Applications (CNNA)

Towards a Mobile Navigation Device

*Kristóf Karacs, Ákos Kusnyerik, Mihály Radványi, Tamás Roska**
Faculty of Information Technology,
Pázmány Péter Catholic University
* MTA-SZTAKI
Budapest, Hungary

Mihály Szuhaj
'IT for the Visually Impaired' Foundation,
Budapest, Hungary

***Abstract*—This paper builds on our recently introduced concept of a portable, non-invasive navigation aid, called a bionic eyeglass, which assists the blind and visually impaired in everyday situations by detecting and recognizing objects and understanding basic events. We present the results of the first online tests carried out by blind subjects on an experimental prototype of the system, which is built on a cellular visual computer. The tests involved laboratory experiments and outdoor live test cases. Users could handle the device easily and the results show that it can provide real help in the given tasks.**

Keywords-navigation aid, cellular computer, visual impairment

I. Introduction

There are numerous situations in which having access to visual information in our local environment is crucial to get around. People with visual impairment cannot do this by themselves; they need the help of another (sighted) person. Although this has positive social aspects, in many cases they would prefer to be independent, or at least have the possibility to choose. The bionic eyeglass could fill this need, integrating many useful functions into a single device. Respected to vision restoration techniques that are actively developed [4], this approach promises to give a solution for end users in the near future, and eliminates any potential health risk involved.

The key challenge in creating a mobile information and navigation aid is to be able to process image flows in real time on a mobile platform and give a feedback within a relatively short time (less than 1-2 sec).

In this paper we analyze the design aspects of such a system and the advantages of the cellular platform that we have chosen respected to other possible realizations. After an overview on the experimental prototype of the system its preliminary evaluation is given through original experiments.

System design aspects, including the analysis of choice of architecture are outlined in Section 2. The experimental prototype is described in Section 3, Section 4 presents the test results, and finally Section 5 concludes the paper.

ONR N00014-07-1-0350
Bólyai János Research Scholarship

II. System design

The term mobile refers to the following requirements in our case: (i) low weight, (ii) small form factor, (iii) autonomy of at least 6-8 hours; whereas low response time depends on two factors: high computing power and low latency.

When faced with the contradictory requirements of having low weight, high computing power, low latency, and long autonomy, as well as being small, anytime accessible, and possibly independent of external infrastructure, we have found only one possibility to satisfy all of them: the use of an embedded cellular platform which is not only mobile, but features a very high speed/power ratio. Systems relying on remote servers can achieve comparably good speed and speed/power values, but they may introduce a latency that is unacceptable for some tasks. However, this does not mean that occasional network connections are ruled out in our system. Indeed, data downloaded from public and personal databases, such as real time geo-location and traffic information could significantly enhance detection and recognition rates for many of the functions, but these operations can take place asynchronously.

The independence from external infrastructure has the following advantages:

- Cheaper operation (no need for a mobile Internet plan)
- Usable at any place, independent of coverage
- Immediate usability everywhere, no need to install markers or any other hardware

To provide input and output interface to the cellular computer a convenient choice is to use a cellular phone, since it is already in the pocket of most people and has adequate input and output modalities: a camera, a microphone and a loudspeaker. In addition, it may have other sensors, such as accelerometer, gyroscope, and GPS.

978-1-4244-6678-8/10/$26.00 ©2010 IEEE

III. Experimental Prototype

The present prototype of the Bionic Eyeglass [3] (see Fig. 1) was built using the Bi-i visual computer [5] as its main computational platform, which is based on the Cellular Neural/nonlinear Network – Universal Machine (CNN-UM) [6, 7] and the underlying Cellular Wave Computing principle. The prototype includes a Wi-Fi enabled cell phone with a built camera that we use as frontend and a wireless adapter connected to the Bi-i computer.

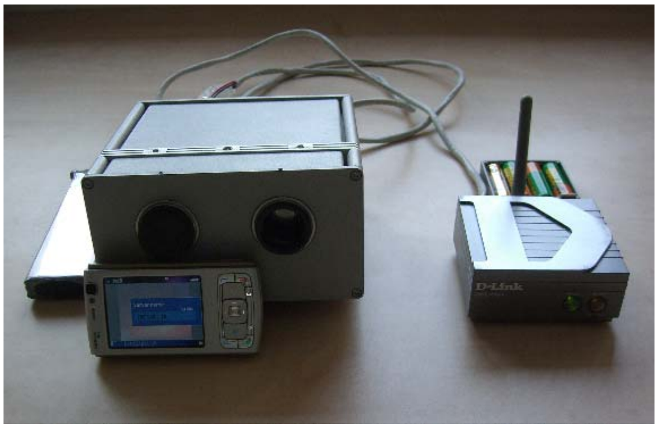

Figure 1. The experimental prototype of the bionic eyeglass. The two way communication between the cell phone and the Bi-i computer is routed through a wireless access point which is connected to the Bi-i with an ethernet cable.

The typical data flow of the system is as follows. The user selects a desired functionality with a numeric key from the cell phone, which activates an algorithm in the Bi-i. The user holds (and moves) the cell phone in the desired direction while the cell phone streams the input video down to the Bi-i. The Bi-i processes the incoming frames with two types of cellular wave computing algorithms: (i) standard templates and subroutines (analogic channels) and (ii) bio-inspired neuromorphic spatial-temporal event detection. Most functions include type i algorithms, such as banknote recognition, crosswalk detection, sign detection and recognition. The second type of algorithm is a neuromorphic saliency system [8] using the recently developed multi-channel mammalian retinal model [9] followed by a classifier using the semantic embedding principle.

Given the final detection or recognition result the encoded decision is sent back to the cell phone, which communicates it to the user via an audio snippet. This can be either a pre-recorded audio file or a generated sound. The system architecture is shown in Fig. 2.

We have previously consulted representatives of visually impaired associations to determine the set of tasks in which the device could bring potential benefit. [1] The functions chosen are summarized in Table I, grouped according to location types they can be used in, with the functions already implemented being italicized. We continuously liaise with potential users through different organizations to keep this list updated.

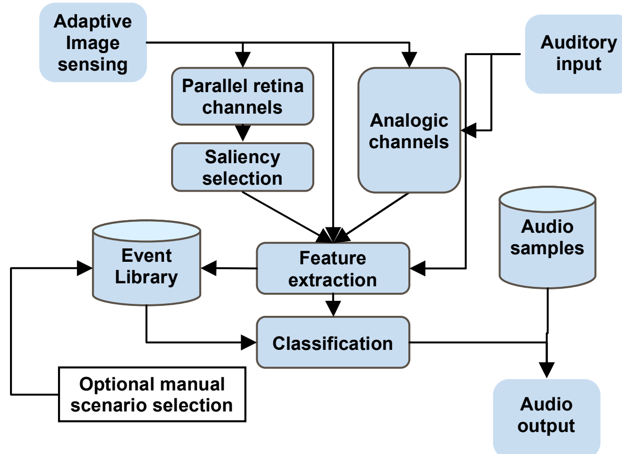

Figure 2. Overview of the system architecture

Table I Planned functionality for the Bionic Eyeglass

<table>
<tr><th></th><th>Home</th><th>Street</th><th>Office</th></tr>
<tr><td rowspan="7">User-initiated functions</td><td>Color and pattern recognition</td><td>Recognition of pedestrian crosswalks</td><td>Button and display recognition in elevators</td></tr>
<tr><td>Bank note recognition</td><td>Public transport: stop identification; sign recognition</td><td>Navigation in public offices and restrooms</td></tr>
<tr><td></td><td colspan="2">Pictogram recognition</td></tr>
<tr><td></td><td colspan="2">Detection and recognition of high contrast client displays</td></tr>
<tr><td></td><td colspan="2">Recognition of messages on screens</td></tr>
<tr><td colspan="3">Motion direction recognition</td></tr>
<tr><td colspan="3">Light source detection</td></tr>
<tr><td>Autonomous warnings</td><td></td><td colspan="2">Obstacles at head and chest level (i.e. boughs, signs, devices attached to the wall)</td></tr>
</table>

IV. Testing

A. *Off-line tests*

Off-line tests are carried out on video flows from a database containing video records taken in situations mentioned in Table I. The database is being continuously expanded, presently there are more than 500 video flows recorded by blind people in it, with lengths between 10 and 90 seconds. Indoor and outdoor recordings were taken under different light conditions.

Commercial cellular phones and digital cameras have been used to record videos. The resolution of the videos is either QCIF or QVGA. Presently no visual microprocessors are available with a resolution higher than QCIF, these recordings were taken for performance comparison purposes. Phones appropriate for this task must have a camera capable of video recording with at least QCIF resolution, and there must be a hard-button by which recording can be started and stopped (soft buttons on a touch-screen are too vague for a visually impaired user).

B. On-line tests

In addition to off-line tests, that are useful for preliminary testing and can used from the early stages of algorithm development, we also test the functions interactively, performing both at-the-scene experiments and standardized laboratory tests. Before performing any of the tests detailed explanation is provided to the volunteer/patient regarding the experiment. All of the measurements are planned in accordance with the Declaration of Helsinki (developed by WMA) ethical principles.

Every volunteer is asked to sign a written informed consent in advance. Each of the tests focus on one specific function in which the user has to detect or recognize targets repeatedly. A short introduction is given to the user on how to hold the camera and what is the maximum speed the camera should be moved with. During the first few trials oral feedback is given on whether the positioning is correct regarding horizontal and vertical axes. Evaluations of the results are always performed after each test.

All features are first tested in a standardized environment setup to provide maximal benefit of the reproducibility and reliability. Following the laboratory tests we also do live tests in sample environments. These tests simulate the conditions of real-life usage, however, always under control.

1) Experiment 1

Detection of different light sources was the purpose of this experiment, which took place in a laboratory environment. Four different light sources were set as targets: a candle light (which is used as a standard measure in ophthalmologic tests), natural light coming through a window, a 17” LCD screen with a document opened full-screen, and a 40W light bulb. The window was located at a distance of 3-5 meters from the subject while the other targets were placed at a distance of 1-2 meters. Detection of a target was signaled by a beeping sound.

The subject was asked to try to locate any targets nearby, given an approximate direction. Results showed that the light bulb and the computer screen could be detected very robustly, i.e., continuous beeping was produced. The window could be detected in most of the cases, but the detection was only reliable when the window moved into the field of view, after 1-2 seconds the detection became intermittent due to the auto-gain feature of the camera. The light of the candle could not be reliably detected from the above mentioned distances; continuous detection was set off only if we moved the candle as close as 50cm to the cell phone.

2) Experiment 2

In this case color recognition was tested in a laboratory environment. Objects of different colors and shades were placed on a table and some clothes hung on a wall, and the subject was asked to try to locate one of the objects and move the camera close to it. Fig. 3 shows some of the objects tested. The recognition could be triggered by the press of a button. After identification of the main colors in the field of view, the system classified them into one of the following classes: black, white, gray, blue, green, red, yellow, brown, and light-blue. Finally the system named at most three dominant ones, and the user, if confident in the result, should have repeated it.

Figure 3. Sample objects tested in Experiment 2.

Based on the results collected we can conclude that (i) for objects of a single color the system could identify the colors with the best accuracy, if the camera was located at a distance of 30–60cm; (ii) ambiguous colors were classified into one of the neighboring classes;(iii) in case of multi-colored objects some of the colors were omitted if the object was not centrally aligned, however the subject did not realize that some colors were missing; (iv) errors were mainly introduced by lighting variances such as reflecting surfaces or strong shadows in the wrinkles of clothes.

3) Experiment 3

The goal in this test was to detect pedestrian crosswalks in real-life situations using the algorithms published in [11]. (The algorithms have been previously tested in a laboratory environment by projecting different test frames from the off-line database on LCD screens, and moving the camera around.) The experiment took place at a big intersection with some divided roads, some roads with tram rails, and some with crosswalks forming an angle of 45° with them. A sample measurement scene is shown in Fig. 4.

At each scene the subject was asked to locate nearby crosswalks as if he/she wanted to cross the street. The subject was given no hint on possible directions of crosswalks around him/her, the only information available was the sound of bypassing traffic. The subject generally had to start the search from a point 5 meter away from the crosswalk, he/she had to approach it and then signal with his arm the most confident direction.

Figure 4. One of the measurement scenes of Experiment 3.

Suboptimal weather conditions lowered the recognition rate for individual frames but even this way, overall video recognition results were good in the sense that all crosswalks were correctly identified and the subject provided a feedback in less than 2 seconds in 80% of the cases. The painted stripes could be successfully localized even if the color filtering resulted in a heavily noisy image (See Fig. 5).

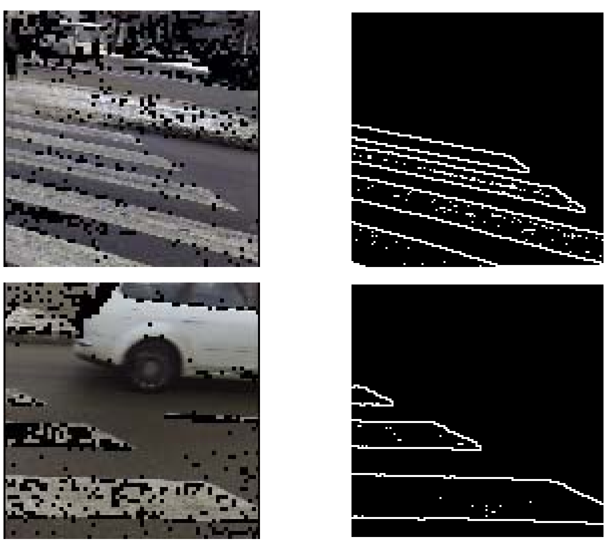

Figure 5. Two processing steps extracted from two frames in Experiment 3. In the left column the output of the gray bandpass filter is shown, whereas the right column contains the results of the edge detection on the located stripes.

V. CONCLUSIONS

As retinal implants are not yet available, an external non-invasive bionic device could be of real use for blind and visually impaired people. Cellular architectures are perfectly suitable for a compact bionic device, enabling real time processing of video flows.

Initial tests of three separate functions on the experimental prototype device have proven the feasibility of the bionic eyeglass concept. The results of the test indicate that in practical cases the functions operate as expected, but detailed analysis could reveal individual in which noise introduces some deviation at some point along the processing.

We plan to implement and test all the other tasks listed in Table I, so that the prototype could have a showcase of applications.

ACKNOWLEDGMENT

The support of the Pázmány Péter Catholic University, the Office of Naval Research as well as the Szentágothai Knowledge Center are kindly acknowledged. We thank Norbert Bérci for his kind help.

REFERENCES

[1] T. Roska, D. Bálya, A. Lázár, K. Karacs and R. Wagner, "System Aspects of a Bionic Eyeglass," in *Proc. of the 2006 IEEE International Symposium on Circuits and Systems (ISCAS 2006)*, Island of Kos, Greece, May 21–24, 2006, pp. 161–164.

[2] K. Karacs, A. Lázár, R. Wagner, D. Bálya, T. Roska, M. Szuhaj, "Bionic Eyeglass: an Audio Guide for Visually Impaired," in *Proc. of the First IEEE Biomedical Circuits and Systems Conference,* London, UK, Dec. 2006, pp. 190–193.

[3] K. Karacs, A. Lázár, R. Wagner, B. Bálint, T. Roska, and M. Szuhaj, "Bionic Eyeglass: The First Prototype, A Personal Navigation Device for Visually Impaired," in *Proc. of First International Symposium on Applied Sciences in Biomedical and Communication Technologies* (ISABEL 2008), Aalborg, Denmark, 2008.

[4] E. Margalit et al., "Retinal Prosthesis for the Blind," *Survey of Ophthalmology*, vol. 47, no. 4, pp. 335–356, 2002.

[5] Á. Zarándy and Cs. Rekeczky, "Bi-i: a standalone ultra high speed cellular vision system," *IEEE Circuits and Systems Magazine*, 2005. vol. 5, no. 2, pp. 36–45.

[6] T. Roska and L. O. Chua, "The CNN Universal Machine: An Analogic Array Computer," *IEEE Trans. on Circuits and Systems – II: Analog and Digital Signal Processing*, vol. 40, pp. 163–173, 1993.

[7] L. O. Chua and T. Roska, *Cellular Neural Networks and Visual Computing*, Cambridge University Press, Cambridge, UK, 2002.

[8] L. Itti, Modeling Primate Visual Attention, in: *Computational Neuroscience: A Comprehensive Approach,* (J. Feng Ed.), pp. 635-655, Boca Raton: CRC Press, 2003.

[9] D. Bálya, B. Roska, T. Roska, F. S. Werblin, "A CNN Framework for Modeling Parallel Processing in a Mammalian Retina," *Int'l Journal on Circuit Theory and Applications*, vol. 30, pp. 363-393, 2002.

[10] L. Kék, K. Karacs, and T. Roska, Eds., *Cellular Wave Computing Library (Templates, Algorithms and Programs) V. 2.1*. MTA-SZTAKI, Budapest, Hungary, 2007.

[11] M. Radvanyi, G. E. Pazienza, and K. Karacs, "Crosswalk Recognition through CNNs for the Bionic Camera: Manual vs. Automatic Design," in *Proc. of the European Conference on Circuit Theory and Design* (ECCTD'09), Antalya, Turkey, Aug. 2009.