Computer Vision-based System for Impaired Human Vision Compensation

Michael Shell
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332–0250
Email: . . .

Homer Simpson
Twentieth Century Fox
Springfield, USA
Email: homer@thesimpsons.com

Andrius Budrionis
Vilnius Gediminas Technical University
Vilnius, Lithuania
Norwegian Centre for E-health Research
University Hospital of North Norway
Troms, Norway

***Abstract*—The abstract goes here. The abstract goes here. The abstract goes here. The abstract goes here. The abstract goes here. The abstract goes here.**

***Index Terms*—Computer vision; Mobile application; Travel aid; Tactile.**

I. Introduction

More than 250 million people worldwide have moderate to severe vision impairment, while ($\approx$ 36 million are blind) [2]. During the past decades significant effort was devoted to develop computer vision and other sensor based aids (e.g. [3], [4], [16], [7]) for helping the blind and visually impaired users to perceive the surrounding world. However, computer vision is rapidly evolving field, and systems based on the aforementioned approaches often lack accuracy and reliability in real-world conditions. In this article we describe a system, which employs modern computer vision techniques for compensation of lost or impaired vision function in humans. Many of the previously proposed electronic aids for the blind count on highly specialized hardware, for instance smart glasses [7], Microsoft Kinnect sensors [14], helmet-mounted photo sensors and cameras [15]. Such advanced and costly technology brings additional complexity into daily tasks, which visually impaired people currently manage by the help of a white cane. This project puts major emphasis on developing a tool, which seamlessly integrates with the hardware visually impaired person already has on-hand and is used to (smartphone). This paper describes a high level architecture of the proposed technology aid for visually impaired people.

II. Method

Following the principles of participatory design [12], [13] we included the final users into the design process of the system. Two visually impaired persons participated in discussion and focus group meetings together with the project team to identify the key properties of the solution. During the six discussion sessions major attention was paid to:

1) Functionality and key features of the solution addressing real-life scenarios of visually impaired persons;
2) Technical feasibility, connecting user needs to the latest development in computer vision and assisted technologies;
3) Interface between the visually impaired person and the assisted technology;
4) Appearance, usability and potential costs of the final product.

Design process involved testing of the existing mobile phone app (Seeing AI) for assisting visually impaired persons. Semi-structured tests were performed by the visually impaired participants accompanied by a researcher from a project group. Strengths and weaknesses of the existing solutions are reflected in the design of the proposed system.

III. Results

User-centered design sessions with visually impaired persons provided substantial input for the design of the proposed system (Fig. 1). The process made it clear that understanding of the user needs was limited in the project group. Based on the results of the user-centered design process, we present the outline of the system for impaired human vision compensation.

The key part of the proposed system is a smartphone device, connecting user interface and data processing components into a well-functioning ecosystem. Smartphone was selected due to its high availability and capability to perform required computations or forward visual information to the external server via the Internet connection. The cornerstone of the proposed solution is a wearable forehead belt, integrating two cameras and a tactile feedback device. Bone conductive headphones are suggested to facilitate audio feedback.

Discussions on functionality and key features of the solution revealed a gap between the aims of the existing tools, which are often focused on advanced navigation and scene description features, and actual user needs. The end users highlighted that a simplistic solution for navigating is lacking. They were not excited about advanced object detection and scene description techniques, which may overcomplicate their daily life. Major focus was put on a easy to use and minimalistic tool, which can be trusted. Direction and approximate distance to obstacle were highlighted as major requirements. To our surprise, only 5-8 distinctive object classes (for instance, buss

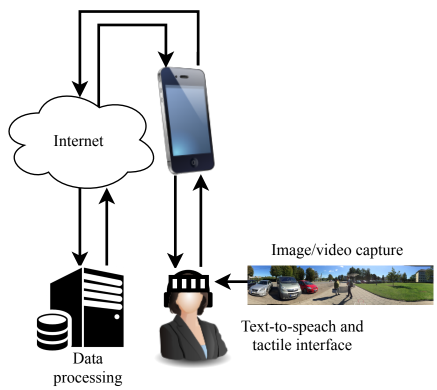

Fig. 1. Schematics of FRP device.

stop, pedestrian crossing, doors, stairs, etc) were of major importance while navigating. Additional information, such as type of a passing car, blossoming flowers in the park or person passing riding a bicycle was considered as overwhelming and distractive.

Social networking functionality was accepted as a potentially important feature allowing other users/volunteers to describe certain objects (for instance, temporary road works) with exact location and description. Navigating around such obstacles can be challenging for computer vision-based tools and may decrease user's trust in the tool. Integration of additional information from other users, governmental organizations and volunteers could be of high value for the visually impaired people.

Computer vision technology is under rapid development and during last years made major breakthrough in performance and efficiency. Deep learning neural networks, which are de facto standard in modern image/video processing, are now well supported by software libraries and dedicated hardware built-in even in relatively low-power computational devices, such as smartphones. These devices are equipped with high-quality cameras, 4G Internet connectivity and sufficient computational power to perform initial data processing locally. The maturity of the aforementioned technologies suggest that our proposed computational aid for visually impaired persons may be highly feasible.

The cornerstone of this project is the interface between the visually impaired person and the assistive technology. We are proposing a combination of audio and tactile feedback, which may improve the interaction between the assitive technology and the user. Similar interfaces were suggested by research communities earlier [7], [16], however, current knowledge on user preferences and evaluation of various options of tactile feedback (placement on the body, frequencies, strength, etc.) is limited.

Appearance, usability and cost of the final solution is of major importance to the users. The proposed assistive technology should replace/supplement the tool that visually impaired persons used for decades - the white cane. Using a white cane requires little to none training, has very low costs and is relatively reliable. It provides information on the surrounding objects 1-2 meters in front of the user. While the cane can used to detect a nearby obstacle, electronic travel aid could be beneficial for longer range (10-20m) route planning and object recognition. However, it should integrate seamlessly into the existing navigation practices of visually impaired people.

IV. Discussion

In this article we outlined an idea of complex computer vision-based sensor architecture, which can help to partially compensate impaired or lost human sight. Main improvements comparing to similar previous work ([**?**], [**?**]) is: ability to use efficient (but still computationally intensive) modern computer vision algorithms via Internet connection, present the signal to the user by combined bone conductive headphone-tactile actuator, low cost and consequently increased potential availability of the device. Suggested architecture can be used as basis for various computer vision-based systems, allowing to help visually impaired users in everyday routines (e.g. outdoor/indoor navigation,...). On the other side, due to remotelly carried visual analysis, in some cases the suggested system may suffer from too large latency, its reliability depends on stability and quality of Internet connection. Such a problem partially may be solved in near future, when 5G networks will prevail.

References

[1] Amos, B., Ludwiczuk, B., and Satyanarayanan, M. Openface: A general-purpose face recognition library with mobile applications. Technical report, CMU-CS-16-118, CMU School of Computer Science, 2016.

[2] Bourne RRA, Flaxman SR, Braithwaite T, Cicinelli MV, Das A, Jonas JB, et al.; Vision Loss Expert Group. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Glob Health. 2017 Sep;5(9):e888-97.

[3] Caraiman, S., Morar, A., Owczarek, M.,Burlacu, A., Rzeszotarski, D., Botezatu, N., Herghelegiu, P., Moldoveanu, F., Strumillo, P., Moldoveanu, A. Computer Vision for the Visually Impaired: the Sound of Vision System. IEEE International Conference on Computer Vision Workshops, pp. 1480-1489, 2017.

[4] Csapó, A., Wersényi, G., Nagy, H., Stockman, T. A survey of assistive technologies and applications for blind users on mobile platforms: a review and foundation for research. Journal on Multimodal User Interfaces. Vol. 9, issue 4, pp. 275-286, 2015.

[5] Dai, J., Li, Y., He, K., Sun J. R-FCN: Object Detection via Region-based Fully Convolutional Networks. In Advances in neural information processing systems, pp. 379-387, 2016.

[6] Poggi, M., Mattoccia, S. A wearable mobility aid for the visually impaired based on embedded 3d vision and deep learning. Proceeding of IEEE Symposium on Compututers and Communication, pp. 208-213, 2016.

[7] Zientara, P.,A., Lee, S., Smith, G., H., Brenner, R., Itti, L., Rosson M., B., Carroll, J., M., Irick K., M., Narayanan, V. Third Eye: A shopping assistant for the visually impaired. Computer Vol. 50, Issue 2, pp. 16-24, 2017.

[8] Liu, Ch., Mao, J., Sha, F., Yuille, A. Attention Correctness in Neural Image Captioning. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, pp. 4176-4182, 2017.

[9] Laina, I., Rupprecht, C., Belagiannis, V., Tombari, F., Navab, N. Deeper depth prediction with fully convolutional residual networks. Fourth International Conference on 3D Vision (3DV), pp. 239-248, 2016.

[10] Ohn-Bar, E., Kitani, K., Asakawa, Ch. Personalized Dynamics Models for Adaptive Assistive Navigation Interfaces. arXiv:1804.04118, [cs.LG], 2018.

[11] Ren, Sh., He, K., Girshick, R., and Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems, pp. 91-99, 2015.
[12] Kensing, F., Simonsen, J., and Bødker, K. MUST: A Method for Participatory Design. Hum-Comput Interact, vol. 13, no. 2, pp. 167-198, Jun. 1998.
[13] Carroll, J. M., Rosson, M. B. Participatory design in community informatics. Des. Stud., vol. 28, no. 3, pp. 243-261, May 2007.
[14] Owayjan, M., Hayek, A., Nassrallah, H., Eldor, M. Smart Assistive Navigation System for Blind and Visually Impaired Individuals. 2015 International Conference on Advances in Biomedical Engineering (ICABME), 2015.
[15] Dunai, L. D., Lengua, I. L., Tortajada, I., Brusola Simon, F. Obstacle detectors for visually impaired people. 2014 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), 2014.
[16] Poggi, M., Mattoccia, S. A wearable mobility aid for the visually impaired based on embedded 3D vision and deep learning. 2016 IEEE Symposium on Computers and Communication (ISCC), 2016