2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)

Sensory Substitution of Vision: A Systematic Mapping and a Deep Learning Object Detection Proposition

Elze P. Lima N., Ronaldo M. Costa, Deborah S. A. Fernandes
Instituto de Informática
Universidade Federal de Goiás
Goiânia - Goiás - Brasil
$\{elzeneto, ronaldocosta, deborah\}$@inf.ufg.br

Fabrízzio A. A. M. N. Soares
Department of Computer Science
Southern Oregon University
Oregon - USA
fabrizzio.soares@gmail.com

***Abstract*—Since 1946 methods for sensory substitution of vision has been studied; however, half a century after the beginning of this line of research, this keep been a massive problem in a world with about 50.6 million people with irreversible blindness. This research presents how self-help devices for visually impaired are approach in recent years and proposes a new approach based on object recognition with deep learning. Through it, it is possible to perceive the trends in this line of research, how devices obtain information from the environment, how they interact with users, and other aspects — pointing essential factors to all those who research or wish to study this area.**

***Index Terms*—Vision Substitution, Computer Vision, Assistive Technologies, Visual Impairment, Object Detection, Deep Learning, Tensor Flow Lite, Mobile.**

I. Introduction

According to the World Health Organization (WHO), near of 253 million people in the world live with a kind of visual impairment, of that, 36 million are blind, and 217 million have a moderate to severe visual loss. Besides that, one fundamental data is that 80% of all visual impairment in the world would not exist with prevention and treatment [1]. For these people, a technological approach is irrelevant because better health conditions would be enough to restore their sight.

The target audience of technologies for sensory substitution of vision is almost 50.6 million people who have a cureless visual impairment, of those 1.4 million are kids with less than 15 years. This line of research could change the way millions of people interact with the world, being an opportunity for independence and social inclusion [2]. Researchers point out that this group would be three times bigger by the years 2050 because of population growth and aging [3].

Even with all needs showed, tools and techniques for sensory substitution of vision have low acceptance by users, and most of them are not adapted for the use of children [4]. Researchers and companies are developing devices along decades; Papers by Haskins Lab presents the use of ultrasonic sensors and others alternatives since 1946 [5], but they were never applied widespread because of long and hard training need [6].

There are many approaches to the development of a device to help visually impaired people; they differ from each other, mainly by hardware, way of sensing, and user interaction [7]. Taking into account the relevance of this problem, the spread of this research line and the number of devices already produced, this paper presents a literature review of this area and gives a proposition as a lead to a good development approach. Making it easier to research groups not execute the same mistakes from decades ago and finally produces acceptable devices and solutions for widespread use.

Following sections show each aspect of a literature review and each step using the software StArt as a model. This project was developed also using review papers as a reference [8] [9] but looking for a robust pipeline. The paper method chosen for literature review was a systematic mapping, and it differs a little from a systematic review; both are secondary studies, systematic mapping has the objective of link related researches by using essential questions, while systematic review evaluates researches qualitatively by research questions.

The systematic mapping starts with planning step found in Section II, in that section the research protocol is followed using the reference software mentioned. After that, following this protocol, papers are collected from the databases. Authors used titles, abstracts, and keywords for selection and after that, read and classified the full texts, shown in Section III. After extraction, Section IV exposes data analyses and Section V and VI a proposition and conclusions.

II. Literature Review Protocol

Sensory substitution of vision is a vast research area, so an excellent way to do a literature review is by a systematic mapping. The steps of a systematic mapping are research questions, protocol, analysis, and conclusion. Analysis and conclusion steps answer relevant questions and objectives. The protocol makes possible repeatability, and the next sections show its steps [8].

2375-0197/19/$31.00 ©2019 IEEE
DOI 10.1109/ICTAI.2019.00274

A. Research Objective

Investigate original studies in the literature that are related to interaction methods for devices used as a sensory substitution of vision. Separate the way that these devices are used to the navigation and environment identification. Then shows the differences between these devices and identify the trends in research lines for this study area.

B. Main Questions

The following questions were relevant using the research objective as a direction:

- What are the methods and techniques used for sensory substitution of vision?
- What is the hardware used in the devices?
- How do interaction and feedback work ?
- What environment these devices are planned to use?

With these questions, it is possible to track tendencies and evolution in this area of research. Knowing which methods are used and techniques along the years in each aspect (hardware, feedback, and environment) it is possible to know the evolution and conclusions already made by researchers.

C. Population

The population target were original papers about devices for sensory substitution of vision. It also included papers focused on interaction with blind users and about in developing devices. Another target is that the papers should be available at high confidence databases [10].

D. Intervention

Evaluation tools and techniques were used as an approach for this paper, systematically showing the adopted metrics.

E. Research Method and Used Strings

Authors chose digital peer-reviewed databases to develop the review. The criteria for selection were the relevance of database for the research area, the regularity of updates, and the availability of bibliography data [10]. Research strings were used to return papers to answer the main questions. Authors choose strings that joined "visual impairment", "sensory substitution", and "guidance". These strings included synonymous and vocabulary variations to get a maximum number of papers. For each database authors used the following strings:

- **Scopus:** *(("blind" OR "visual impairment" OR "visually impaired") AND ("sensory substitution" OR "navigation device" OR "Vision substitution" OR "Self-Help Devices" OR "Assistive technologies") AND (navigation OR guiding OR Guidance OR wayfinding))*;
- **ScienceDirect:** *(("blindness" OR "visually impaired") AND ("sensory substitution" OR "navigation device" OR "Vision substitution" OR "Self-Help Devices" OR "Assistive technology") AND (navigation OR guiding))*;
- **IEEE:** *((("blind" OR "visual impairment" OR "visually impaired") AND ("sensory substitution" OR "navigation device" OR "Vision substitution" OR "Self-Help Devices" OR "Assistive technologies") AND (navigation OR guiding OR guidance OR wayfinding)))*;
- **PubMed:** *((("blind" OR "visual impairment" OR "visually impaired")) AND ("sensory substitution" OR "navigation device" OR "Vision substitution" OR "Self-Help Devices" OR "Assistive technologies")) AND (navigation OR guiding OR guidance OR wayfinding)*.

F. Control

The control group was composed of five articles [11] [12] [13] [14] [15]. These papers have the most number of citations in this area of research and present fully developed devices. They indicated if a string was returning expected results by their show or not in the research. If a string was not including one of the five notorious articles, it should be changed until they were getting in the research.

G. Results

The authors wanted to obtain a broad vision and knowledge of the diversity of methods and techniques used to a sensory substitution of vision — also an identification of trends in this research line.

H. Application

Researchers in the research line of sensory substitution of vision that develops devices and tools for interaction with visually impaired people can use this systematic mapping. This paper identifies opportunities for research, challenges in their development and trends in this line of search.

I. Tools and Instrumentation

The databases used in the research were: Scopus, ScienceDirect, IEEE Xplore, and PubMed. They are regularly updated, and the first three are essential in the area of computer science and the last one in the medical area, this is also important in the research line of sensory substitution of vision [10].

After getting research results using previous strings, papers were uploaded to StArt software. The software has three steps: planning, execution, and summarization. Planning presents and builds the review protocol, execution step identifies, select, and preliminary extract data from relevant papers and summarization presents conclusions and results [9].

J. Study Selection Criteria

Primary criteria used to obtain results in each search session for different databases:

- Papers published in English;
- Papers on indexed databases.

Authors choose these criteria to help others researches to reproduce the review. After research using initial criteria, the authors used other rules for selection and extraction. The criteria for selection inclusion were:

- Studies about the interaction of visually impaired people with assistive technologies;
- Devices and prototypes for sensory substitution of vision;

The authors choose these criteria to select the papers about developing a device for help in the sensorial substitution of vision. The exclusion criteria were:

- Sensory substitution of others senses besides vision;
- Posters and short papers researches;
- Do not identify objects or help in navigation;
- Papers about the same device;
- Articles exclusive about medical aspects without use of devices or software;
- Researches published before 2013;
- Researches that are only available in books.

With these exclusion criteria, authors excluded papers about other senses, papers with only medical aspects, and papers not relevant for the research of this kind of devices.

The authors used different categories and their expressions for data extraction. These data made it possible to obtain relevant data for answering the main questions quantitatively. These categories were:

- **Database**. Register database used at each search session: (a) Scopus, (b) ScienceDirect, (c) IEEE Xplore, (d) PubMed;
- **Kind of device**. Register which hardware was used: (a) Smartphone, (b) Developed by author, (c) Smart Glasses, (d) Kinect, (e) Find in the market;
- **Device positioning in user**. Register the place the device is used: (a) Glasses, (b) Cane, (c) Walker, (d) Clothes, (e) Hand, (f) Shoes, (g) Fixed place in room;
- **User interaction**. Register how the feedback is sent to the user: (a) Audible, (b) Tactile in skin, (c) Tactile in tongue, (d) Tactile by vibration in hand;
- **Application**. Register which environment the device was developed for: (a) Indoor, (b) Outdoor, (c) Specific room, (d) General;
- **Sensor used**. Register which sensing method was used: (a) Camera, (b) Infrared, (c) Ultrasonic, (d) GPS, (e) RFID or NFC, (f) Compass or accelerometer, (g) Laser, (h) BLE or WiFI, (i) Other.

III. Systematic Mapping

Following sections will expose the execution and results obtained during the research after established the protocol. After all search sessions, the authors obtained 570 papers, but 152 were duplicates. This massive number of duplicates, 26.7% of all papers, happened because of the use of different sources. Scopus database returned the most number of papers; the authors made de research between May and July of 2018. The Table I will express the number of papers considerate in each step.

A. Selection

The research reviewed 418 papers during the selection step. Each of them evaluated with inclusion and exclusion criteria. Titles, abstracts and keywords were read and considered for acceptance or rejection. Those 212 papers passed the criteria, 50.7% of all papers were selected for the extraction as can be seen in Table I.

B. Extraction

In the extraction step 212 papers were read and classified by inclusion and exclusion criteria. Of 212 papers, only 92 or 43.4% were selected for analysis. From the total of papers, 22% were selected to analysis as seen in Table I.

TABLE I
NUMBER OF PAPERS APPROVED IN EACH STEP

Database	Preliminary Search	Selection	Extraction
ScienceDirect	22	9	8
Scopus	386	177	74
IEEE Xplore	95	22	8
PubMed	67	4	2
Total	**570**	**212**	**92**

IV. Data Analysis

In this section, the data obtained in the systematic mapping was analysed to answer each of the research questions.

What are the methods and techniques used for sensory substitution of vision?

The most used techniques and methods are the ones that use cameras for environment sensing and give audible feedback as seen in Table II and Fig. 2. These methods are the most used because of the level of abstraction and details that can be caught by a camera and processed with computer vision techniques. In this aspect, some tendencies are the use of object recognition, techniques used in autonomous vehicles, and methods for plate, text and people recognition [16] [17].

What is the hardware used in the devices?

Many researchers develop a device instead of using a third party one, but there is also a tendency in using smartphones as a device to help visually impaired people as seen in Fig. 1. The expansion in the use of smartphones all around the world and adaptability for visually impaired people explains the increase in the development in this trend. Another important aspect is the plurality of uses for a smartphone and hardware capabilities in a low price device [18].

How do interaction and feedback work?

In concerning of interaction method for feedback, the most used are audible and tactile by vibration in device hold at hand as seen in Fig. 2. Others methods can be seen as tactile using the tongue to guide, a matrix of motors for vibration in chest and others, but these techniques are invasive and need a long time of adaptation and training [15].

What environment these devices are planned to use?

The development of the majority of devices was for general use, 39 devices for both indoor and outdoor. About specific applications, 20 more devices are for indoor use than outdoor. This difference can be explained by the ease to control and test in indoor scenarios [19].

V. Deep Learning Proposition

After the systematic review, authors get some conclusions for the development of a useful prototype. Some essential aspects of a device for sensorial substitution of vision are:

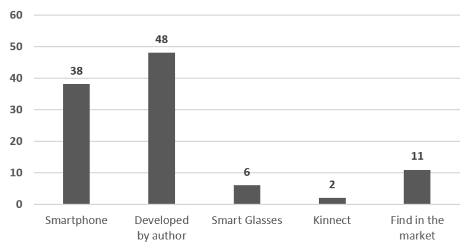

Fig. 1. Kind of device.

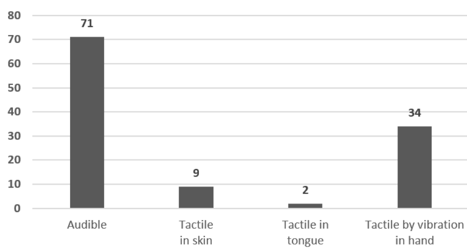

Fig. 2. Device positioning in user.

- The device needs to be easy to use, without training or little training need. Many devices observed in the research have a problematic and extensive training, making the users prefer traditional methods as cane and dogs then learning another difficult one [15].
- The device needs to have a low price, be accessible, and easy to obtain. Solutions with good results may be unavailable to many if they have expensive hardware to be acquired.

Regarding the answers in section IV, authors chose to use a smartphone to develop the proposition. Smartphones cover all tendencies observed by the answers. Authors used the smartphone camera as a sensor that has a high level of information to be worked and also the possibility to use computer vision. Smartphones are a worldwide spread device, have proper battery management, easy to be bough, the possibility of updates from the internet and a lot of already adapted software and other uses for people with visual impairment.

TABLE II
DEVICE POSITIONING AND SENSOR USED BY OCCURRENCE

Device Positioning in User	N[a]	Sensor Used	N[a]
Glasses	20	Camera	48
Cane	11	Infrared	3
Walker	2	Ultrasonic	19
Clothes	21	GPS	16
Hand	43	RFID and NFC	3
Shoe	4	Accelerometer and Compass	18
Environment	5	Laser	1
		BLE and WiFi	5
		Other	9

[a]Number of devices and sensors.

Users can hold a smartphone at hand and clothes, and use it anywhere.

Authors decided not to entirely change the way that a person with visual impairment interacts with the world, but to help in problems that they can not solve without the guidance of other person, like find a small object, a door in a room and others possibilities.

To solve this kind of problem, the authors needed object detection in a smartphone. A right approach for this is by using fast object detection, and for that, authors choose deep learning techniques like Yolo and RCNN [20]. Like seen in Fig. 3 a remote could be detected in 120 milliseconds in a Moto Z2 Force smartphone.

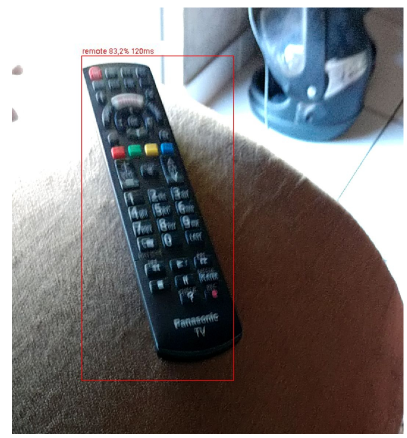

Fig. 3. Test detecting a remote in 120ms.

The framework choose was Tensor Flow Lite. The authors used it to develop an initial software and start to test the system with different techniques. The system detects an object that the user asks to be found by voice recognition, and after that, it guides the user by voice commands and vibration. The user has to move the smartphone, and it starts to vibrate if happens an object detection in the image, after that, the user tries to set the object in the centre of the camera image, using the vibration strength and audible direction commands as a guide.

VI. CONCLUSION AND FUTURE WORKS

The research made was split into steps of a systematic mapping; they are planning, execution and data analysis showing criteria for the steps. These criteria made able validation and reproducibility of this research. It was exposed the protocol construction, data obtained, and the conclusions get. Following the protocol, only 92 papers of 418 were considerate relevant to a systematic mapping of devices for sensory substitution of vision.

From these 92 papers, some relevant observations were getting: most devices use cameras as a sensing method to obtain data from the environment. That happens because cameras make able to obtain complex data as specific object detection, dangerous situations and face detection, making it more accessible to visually impaired people be guided in both static and dynamic environment [21].

Other concerning found as a tendency in the papers were using low cost and multipurpose devices. That is seen by the use of hardware like smartphones and its sensors as a base for developing software and an interactive way for user feedback [18].

From the initial tests of the proposition, the authors notice that using smartphones for developing a device for sensory substitution of vision is a feasible way to help millions of people and that is being possible because of the development of hardware for smartphones and the possibility of using deep learning techniques in small and portable devices. [16] [21] [22]. More tests will be made with a larger group to see the viability of this approach.

Acknowledgment

The authors are grateful for the support of INF-UFG (Instituto de Informática da Universidade Federal de Goiás) and the sponsor of CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior).

References

[1] R. R. A. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg, K. Naidoo, K. Pesudovs, S. Resnikoff, A. Silvester, G. A. Stevens, N. Tahhan, T. Y. Wong, H. R. Taylor, R. Bourne, P. Ackland, A. Arditi, Y. Barkana, B. Bozkurt, T. Braithwaite, A. Bron, D. Budenz, F. Cai, R. Casson, U. Chakravarthy, J. Choi, M. V. Cicinelli, N. Congdon, R. Dana, R. Dandona, L. Dandona, A. Das, I. Dekaris, M. D. Monte, J. Deva, L. Dreer, L. Ellwein, M. Frazier, K. Frick, D. Friedman, J. Furtado, H. Gao, G. Gazzard, R. George, S. Gichuhi, V. Gonzalez, B. Hammond, M. E. Hartnett, M. He, J. Hejtmancik, F. Hirai, J. Huang, A. Ingram, J. Javitt, J. Jonas, C. Joslin, J. Keeffe, J. Kempen, M. Khairallah, R. Khanna, J. Kim, G. Lambrou, V. C. Lansingh, P. Lanzetta, J. Leasher, J. Lim, H. Limburg, K. Mansouri, A. Mathew, A. Morse, B. Munoz, D. Musch, K. Naidoo, V. Nangia, M. Palaiou, M. B. Parodi, F. Y. Pena, K. Pesudovs, T. Peto, H. Quigley, M. Raju, P. Ramulu, S. Resnikoff, A. Robin, L. Rossetti, J. Saaddine, M. Sandar, J. Serle, T. Shen, R. Shetty, P. Sieving, J. C. Silva, A. Silvester, R. S. Sitorus, D. Stambolian, G. Stevens, H. Taylor, J. Tejedor, J. Tielsch, M. Tsilimbaris, J. van Meurs, R. Varma, G. Virgili, J. Volmink, Y. X. Wang, N.-L. Wang, S. West, P. Wiedemann, T. Wong, R. Wormald, and Y. Zheng, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis," *The Lancet Global Health*, vol. 5, no. 9, pp. e888 – e897, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2214109X17302930

[2] S. Sacks, L. Kekelis, and R. Gaylord-Ross, *The development of social skills by blind and visually impaired students: Exploratory studies and strategies*. American Foundation for the Blind, 1992.

[3] S. R. Flaxman, R. R. Bourne, S. Resnikoff, P. Ackland, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen *et al.*, "Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis," *The Lancet Global Health*, vol. 5, no. 12, pp. e1221–e1234, 2017.

[4] M. Gori, G. Cappagli, A. Tonelli, G. Baud-Bovy, and S. Finocchietti, "Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children," *Neuroscience & Biobehavioral Reviews*, vol. 69, pp. 79 – 88, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0149763415302864

[5] L. Kay, "An ultrasonic sensing probe as a mobility aid for the blind," *Ultrasonics*, vol. 2, no. 2, pp. 53–59, 1964.

[6] M. A. Williams, C. Galbraith, S. K. Kane, and A. Hurst, "Just let the cane hit it: how the blind and sighted see navigation differently," in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. ACM, 2014, pp. 217–224.

[7] W. Elmannai and K. Elleithy, "Sensor-based assistive devices for visually-impaired people: current status, challenges, and future directions," *Sensors*, vol. 17, no. 3, p. 565, 2017.

[8] P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain," *Journal of systems and software*, vol. 80, no. 4, pp. 571–583, 2007.

[9] E. Hernandes, A. Zamboni, S. Fabbri, and A. D. Thommazo, "Using gqm and tam to evaluate start-a tool that supports systematic review," *CLEI Electronic Journal*, vol. 15, no. 1, pp. 3–3, 2012.

[10] A. Alnafessah, M. Al-Ammar, S. Al-Hadhrami, A. Al-Salman, and H. Al-Khalifa, "Developing an ultra wideband indoor navigation system for visually impaired people," *International Journal of Distributed Sensor Networks*, vol. 12, no. 7, p. 6152342, 2016.

[11] J. Cecílio, K. Duarte, and P. Furtado, "Blindedroid: An information tracking system for real-time guiding of blind people," *Procedia Computer Science*, vol. 52, pp. 113–120, 2015, cited By 7.

[12] H. Fernandes, J. Faria, P. Martins, H. Paredes, and J. Barroso, "Rfid mesh network as an infrastructure for location based services for the blind," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8008 LNCS, no. PART 5, pp. 39–45, 2013, cited By 2.

[13] S. Al-Khalifa and M. Al-Razgan, "Ebsar: Indoor guidance for the visually impaired," *Computers and Electrical Engineering*, vol. 54, pp. 26–39, 2016, cited By 2.

[14] D. Croce, P. Gallo, D. Garlisi, L. Giarre, S. Mangione, and I. Tinnirello, "Arianna: A smartphone-based navigation system with human in the loop," *2014 22nd Mediterranean Conference on Control and Automation, MED 2014*, pp. 8–13, 2014, cited By 10.

[15] S. Maidenbaum, S. Hanassy, S. Abboud, G. Buchs, D.-R. Chebat, S. Levy-Tzedek, and A. Amedi, "The 'eyecane', a new electronic travel aid for the blind: Technology, behavior & swift learning," *Restorative Neurology and Neuroscience*, vol. 32, no. 6, pp. 813–824, 2014, cited By 33.

[16] R. Tapu, B. Mocanu, and T. Zaharia, "Deep-see: Joint object detection, tracking and recognition with application to visually impaired navigational assistance," *Sensors*, vol. 17, no. 11, p. 2473, 2017.

[17] M. Martinez, A. Roitberg, D. Koester, R. Stiefelhagen, and B. Schauerte, "Using technology developed for autonomous cars to help navigate blind people," in *Computer Vision Workshop (ICCVW), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1424–1432.

[18] A. Rodrigues, K. Montague, H. Nicolau, and T. Guerreiro, "Getting smartphones to talkback: understanding the smartphone adoption process of blind users," in *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. ACM, 2015, pp. 23–32.

[19] A. Alnafessah, M. Al-Ammar, S. Al-Hadhrami, A. Al-Salman, and H. Al-Khalifa, "Developing an ultra wideband indoor navigation system for visually impaired people," *International Journal of Distributed Sensor Networks*, vol. 12, no. 7, p. 6152342, 2016.

[20] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv*, 2018.

[21] R. C. Gonzalez and R. E. Woods, "Image processing," *Digital image processing*, vol. 2, 2007.

[22] A. Arditi and Y. Tian, "User interface preferences in the design of a camera-based navigation and wayfinding aid." *Journal of Visual Impairment & Blindness*, vol. 107, no. 2, 2013.