

## Design of Smart Glasses that Enable Computer Vision for the Improvement of Autonomy of the Visually Impaired.

C. C. Spandonidis\*, D. Spyropoulos, N. Galiatsatos, F. Giannopoulos and D. Karageorgiou

Prisma Electronics S.A., Dimokratias Ave., 68132, Alexandroupoli, Greece

Received 2 March 2021; Accepted 24 May 2021

### Abstract

Recent advances in computer vision brought us closer to the levels of human vision without the need for high computing power. Developments at the algorithmic level are complemented by the wide availability of visual content data necessary for the training of computer vision, as the ever-increasing digitization of society is accompanied by the large-scale production and sharing of digital data. Working on that direction, the e-Vision project aims at the development of an assistive system for people with a reduced vision which enables the use of an array of sensors (camera and headset) mounted on special glasses to transfer the field of vision to their mobile device. The system, in turn, informs the users through the acoustic diode about the object or person in front of them. In this work, the design and development of the sensor system adapted to glasses that will support communication with the application of e-vision via mobile phone are presented. Sensor data are being transmitted in real-time, through different wireless communication approaches. Details regarding initial requirements and feedback from the pilot testing are also provided. It is shown that the system can support its mission for the improvement of the autonomy of the visually impaired.

*Keywords:* e-Vision; Computer Vision; Smart Glasses; Wireless Communication; Data Rate

### 1 Introduction

People with total vision loss on the planet are estimated at 36 million (up from 30.6 million in 1990) [1]. Besides, it is estimated that by 2050 this number will increase to almost 115 million, while the corresponding number of people with limited vision ability is estimated to increase from 217 million today to 588 million in 2050 [2].

According to [3], the global market for assistive technologies for people with limited vision is valued at US\$3,400 million in 2017 and is expected to reach \$7,100 million by 2026, expanding to a CAGR of 8.5% from 2018 to 2026. Advanced technological solutions with a high degree of maturity are expected to encourage healthcare providers, providers of reception and service providers for people with limited vision. Significant investments could be made in developing advanced solutions to address patient needs and improve the operational efficiency of the services available. Table 1 summarizes the segmentation that currently prevails in the market for technological solutions to support people with reduced or no vision [3].

Increasing awareness of assistive technologies for the visually impaired, technological developments, and population growth of target patients are likely to increase the global market from 2018 to 2026. Indicative solutions for object identification are described in [4] and [5], while in [6] a more advanced solution is presented. In [7] and [8] the use of smart glasses in different applications is critically introduced while in [9] and [10] the use of ultra-sound and

acoustic technology for the support of visually impaired is further investigated.

**Table 1.** Technological solutions segmentation

Market Segmentation	
Per product	Per end user category
educational devices & software	blind schools
Mobility devices	enterprises & social organizations
Smart Assist Devices	personal use federation & hospital, and others.

The e-Vision project aimed to exploit recent important results in the field of machine vision, so that systems to assist visually impaired people can now provide the possibility of identifying objects and persons at a semantic level. This work focuses on the design and development of the smart glasses; in the means of an array of sensors (camera and headset) mounted on special glasses, that enable the transfer of users' field of vision to their mobile device and this, in turn, informs them through the acoustic diode about the object or person in front of them.

The rest of the paper is structured as follows: In Section 2 the basic architecture of the e-Vision system is briefly presented. The requirements upon which the design of the sensor array solution relied on is introduced and described in Section 3, where both the initial and the updated requirements introduced after pilot testing have been included. In Section 4, the two solutions that relied on Commercial-off-the-self

\*E-mail address: rdprojects@prismael.com.com

ISSN: 1791-2377 © 2021 School of Science, ITHU. All rights reserved.

doi:10.25103/jestr.142.15

(COTS) components are presented. In Section 5, the design of a combined solution that could benefit from all architectures presented in the previous sections is described on hardware and operation modes level. Finally, in Section 6, the key results of the study are discussed.

## 2 System architecture

Figure 1 presents a general form of the system architecture, in which its building blocks are displayed.

As shown the overall system consists of the following parts:

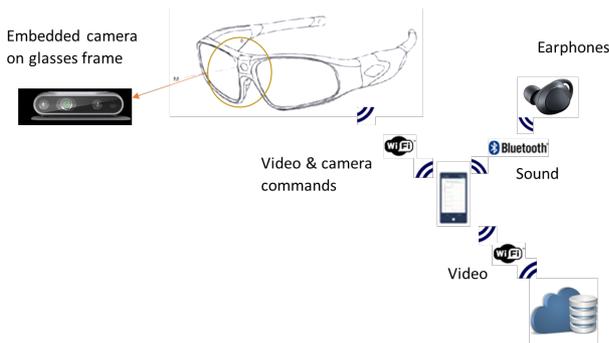


Fig. 1 System Architecture

- Smart glasses that incorporate optical sensors and send data using Wi-Fi technology.
- An audio module integrated to the glasses that supports connectivity via Bluetooth wireless pairing and produces audio effects and information to guide the user.
- Mobile phone app with a choice of services to analyze the image and provide information to headphones.
- Application in a cloud environment to run identification algorithms and transfer the information to the end user.

## 3 Smart glasses requirements

The selection of the smart glasses architecture and their parts was based on three pillars:

- The development of a system covering all the technical specifications regarding the transfer of real-time high-definition video streaming [11].
- The development of a final solution that is economically robust, so that it can result in a commercial product easily acquired by the end user.
- The creation of a device that ensures the discreet support of end users, without creating ergonomic or social exclusion problems, as requested by end users during interviews.

A second criterion was added following the techno-economic analysis of the existing solutions since it showed that the cost of the existing solutions is particularly high. Besides, both the size and layout of the sensors and the final weight of the system create ergonomic problems in the final system, especially when it is carried on the go.

### 3.1 Integration phase

During the integration of the system, i.e., the interconnection of the hardware with the software, the need to diversify the system architecture for downloading an image from a remote server and for sending one to it was discussed within the consortium and the following decisions were reached:

- Send photos instead of high-resolution video to a remote server.
- Examine the ability to connect the sensors directly to the cloud without the mediation of the user's mobile phone.

### 3.2 Pilot testing

The development phase was succeeded by the pilot trials. Participants in the pilot trials were visually impaired people who volunteered to help with the evaluation of the system. A total of nine volunteers participated in the pilot trials (7 men and 2 women with an average age of 32 years). Before pilot trials, Participants participated in an in-house training in the proper use of the system by the qualified technical staff of the project working group and had the opportunity to test the system in a simulation of three scenarios: cultural walk, visit public service, visit the supermarket.

During the pilot trials, the requirement for sending photos in place of high-resolution video were confirmed. Also, the satisfactory operation of the cloud interface architecture adopted by the existing solution was established. However, the requirement for interconnection in an in-line manner has been confirmed, especially in cases of closed space, where the provision of wireless coupling is stable and continuous.

## 4 Commercial-off-the-self solutions

Following the initial requirements, the development of the final system was based on the selection of the appropriate set of high-resolution sensors and cameras, as a single solution with the additional option of wirelessly sending (via Wi-Fi) the high-resolution video to the end user's mobile device. Besides, the selection of a headset for transmitting and playing the audio using the Bluetooth protocol was the most appropriate solution for voice transfer. Commercial-off-the-self products were preferred at this stage both due to their immediate availability, the fact that they could be used to form a prototype quickly and time limitations of the project. Integration and testing of the sensors as well as proof of concept were of great importance at this stage. While the selection of vision-impaired dedicated cameras and/or solutions could be used and were initially investigated, alternative solutions were adopted at the end. The two solutions, that proved to be of great performance, are described below.

### 4.1 Smartphone as Wi-Fi access point

A pair of glasses with a built-in camera and a wireless headset was used. The image from the built-in camera is stored locally on an SD card but is also sent via wireless pairing (Wi-Fi) to the user's device (mobile) for processing. The same pairing allows the camera and communication system to be set up and operated. Figure 2 illustrates the Wi-Fi architecture of this solution.

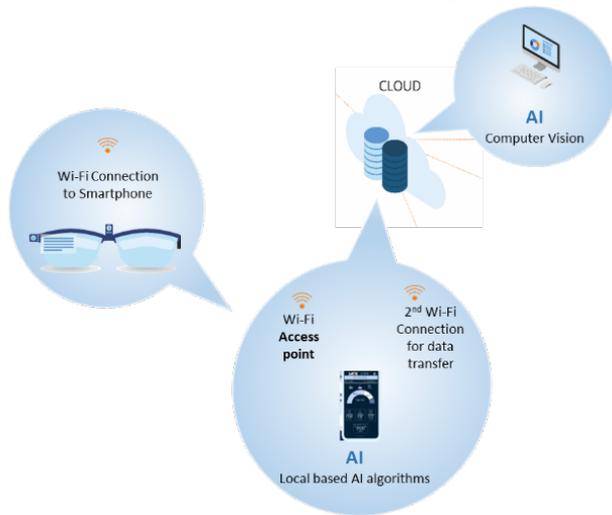


Fig. 2 Wi-Fi architecture of the 1<sup>st</sup> solution.

The image capture and transmission system have been built and implemented so that it can be used several times during the day. It is also designed to maintain its operation for several hours. Its design is flexible and can be adapted to every type of person. Besides, inside it, there is some protective material, specially designed for cooling. It features a high definition 1080P HD video camera lens. It also has built-in Wi-Fi with which live streaming of data can be done. The material of the glasses is designed to be antiallergic, durable, and able to eliminate heat. Its color is matte black. As far as the operating part is concerned, the camera can be adjusted at either 12 degrees up or 10 degrees down, so that it can take the best possible image. The supports on the nose can be adjusted for different sizes that do not make it difficult for the user to use for a long time. The produced system can also control parameters such as the sampling frequency for visual recognition and transmission rate for acoustic communication and change their values according to the user's preferences, ensuring harmonic coexistence with other perceptual processes. The glasses are charged via a USB port, through which the user can also communicate with the pair of glasses in real-time, access the stored images, bypass the need to use the wireless link and/or extend the operating time of the system (e.g. while charging from either the user's mobile phone or a power bank).

Although the solution performed well during the pilot trials and was adopted as the main solution for the e-Vision system, its ability to perform well in real-time high video streams was challenged mainly due to time delays associated with enabling and disabling Wi-Fi via the user's mobile device.

#### 4.2 Direct connection to the Wi-Fi

To satisfy the above requests, an alternative solution was evaluated based on different COTS components. This time the smart glasses were able to directly send images to the cloud. Figure 3 illustrates the W-Fi architecture of this solution.

The used eyeglasses have a built-in camera and microphone on them so that video can be recorded at any time. Their function is based on a fixed Wi-Fi connection, through which the information is sent. The camera's resolution reaches 1280×720 pixels, allowing the device to record high-quality video, either in bright or dark spaces. Also, on this product, there is a built-in slot for a MicroSD card so that the user can save video or photos to it. Video files

and photos are stored on the micro-SD memory card and the user can transfer them to their computer using a USB cable.

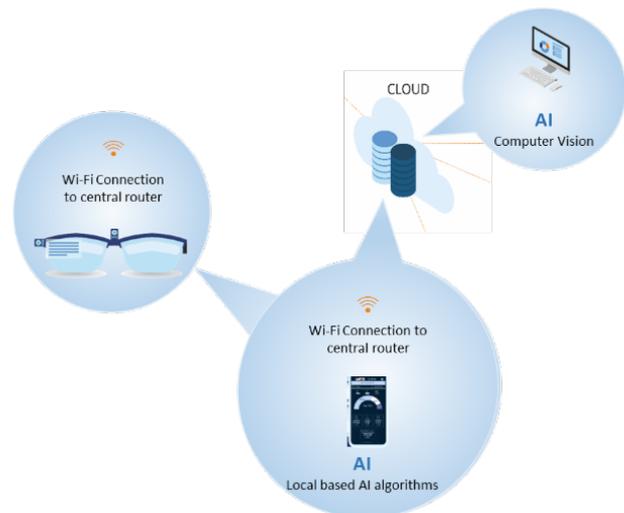


Fig. 3 Wi-Fi architecture of the 2<sup>nd</sup> solution.

Although this alternative meets the additional requirements regarding the direct upload of images to the cloud, it has presented some problems with its exploitation in practice. More specifically, the ability to transfer the image to the cloud allows the system to be exploited only indoors or in areas with free access to the network. In the case of outdoor areas, the images should be sent using the user's mobile device.

## 5 Combined Solution Design

While the use of COTS provided a series of benefits, specific limitations were raised for each architecture. To benefit from both worlds, the authors proceeded with the design and development of a platform that could collect information from the camera installed on the glasses and then send it via Wi-Fi protocol either to the user's mobile phone or to the cloud directly. The platform is based on Prisma-Sense technology which has been used in maritime and smart city applications (e.g. [12], [13]). The operation is selected by the user and the desired mode can be pre-selected. Besides, the circuit presented in this section was created to be integrated into glasses for the visually impaired to identify the objects that the user sees and to return information about them via an audio message.

### 5.1 Hardware

The function of the device is based on the ESP32 micro-controller to which the camera is connected and via the on board Wi-Fi module, the camera image is sent for processing. The description of this image is then returned in audio format, which the user can hear through a speaker. The 3D model of the prototype board is presented in Figure 4.

Esp32-WROOM-32D is a powerful integrated with built-in Wi-Fi + BT + BLE MCU units, that aims at a wide variety of applications, from low-power sensor networks to more demanding ones, such as voice encoding, IoT applications, etc. This chip consists of two cores that can be controlled individually with an adjustable clock frequency, from 80 MHz to 240 MHz. This connection is possible as this controller supports various communication protocols, such as

high-speed SPI, UART, PWM, DAC, I2S, and I2C. The user can continuously monitor the peripherals for their proper operation or any changes. Power to the circuit is powered by a USB cable or battery.

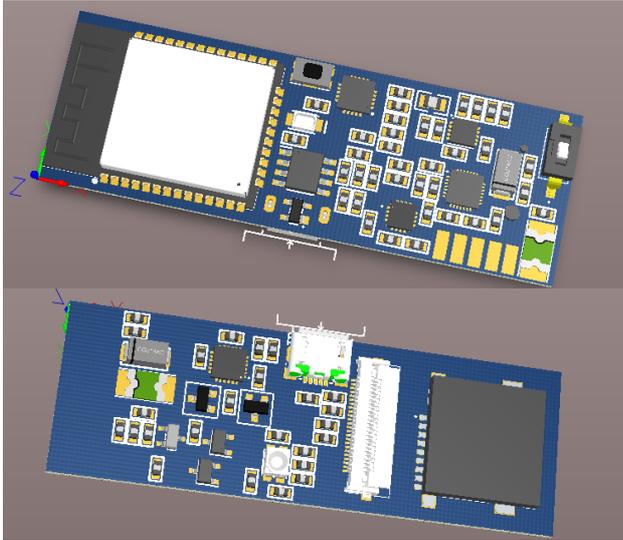


Fig. 4 PCB 3D Top (up) and Bottom (down) Layer

### 5.1.1 Camera operation

The camera used is the OV2640. It is a low-voltage CMOS image sensor that provides the full functionality of a UXGA camera (1632x1232) and an image processor. OV2640 provides 8-bit/10-bit images in various file formats. It is also capable of taking photos at a frequency of up to 15 frames per second (fps) at UXGA resolution that can control image quality, formatting, brightness, color saturation, etc. The schematic depicting the camera module's connection is presented in Figure 5.

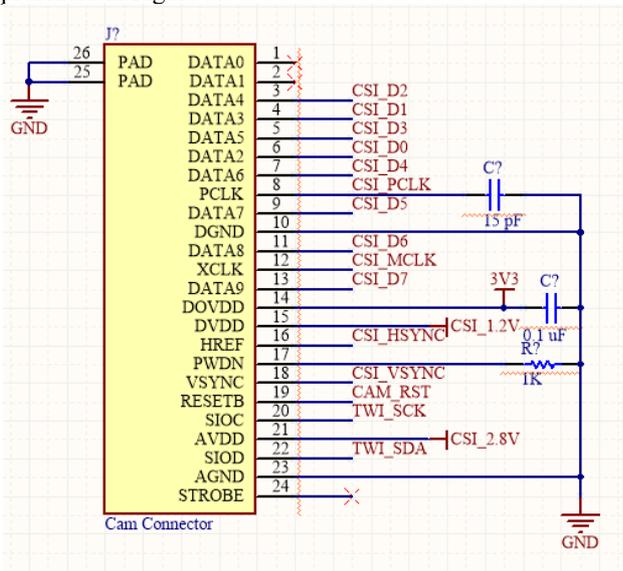


Fig. 5 Camera Schematic Connector

All the above functions are programmable through the Serial Camera Control Bus (SCCB) interface. In order for the system to be able to interact with the camera, it needs to be integrated into an FFC-type connector circuit. Moreover, voltage regulators that can convert the system input voltage, 3.3V, to 2.8V and 1.2V, which are necessary for the operation of the camera are also needed. This camera interface can be

connected directly to most cameras with Digital Video Port (DVP) output, such as OV2640, OV7670, OV7725, etc.

### 5.1.2 Audio decoding

The circuit consists of the integrated MAX98357A. It is an easy-to-use, low-cost Digital Pulse-Code Modulation (PCM) amplifier that delivers Class AB audio performance with Class D efficiency. The digital audio interface automatically recognizes up to 35 different PCM and TDM timing formats, eliminating the need for I2C programming.

The function is further simplified by eliminating the need for an external MCLK signal commonly used for PCM communication. The interface accepts sample rates from 8kHz to 96kHz for all supported data formats. Finally, the integrated circuit can be configured to produce sound on the left channel, on the right channel, or both channels, thus producing stereo output. In the case of this circuit, the sound produced is mono, transmitted via a single channel. The audio module and its connections to the ESP32 microcontroller is depicted in Figure 6.

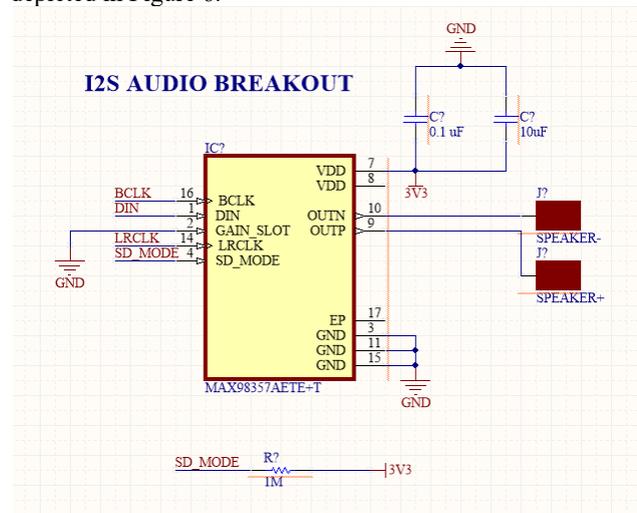


Fig. 6 Audio Decoding Schematic Circuit

### 5.1.3 Memory

ESP32 is connected to an external PSRAM with a capacity of 4Mbyte, increasing the device's RAM and making it possible to download and edit an image. The corresponding integrated circuit used in the layout is IS62WVS5128GBLL. This is a fast 4Mbit-serial static RAM, organized in 512K bytes by 8 bits. The device is accessed through the Serial Peripheral Interface (SPI). ESP32 is properly connected to a MicroSD card slot so that the user can save the photos they take and the audio messages they receive.

## 5.2 Operation modes

In this sub-section, the various operation modes of the device are presented, following the sequence that the user would normally follow in order to operate the device. The process begins with the powering of the device, giving the opportunity to operate in two modes. The first one allows the board to be used as a Wi-Fi access point, while the second one enables the connection to an existing Wi-Fi network, via a dedicated page that allows the user to enter the network credentials (network name, password).

### 5.2.1 1<sup>st</sup> mode– Access point

A computer, smartphone, or tablet, located near the device can connect without a password on the available Wi-Fi network "EVISION" that has been created. After connecting to it, the menu presented in Figure 7 will automatically open.

In case the device is connected to a serial monitor for debugging purposes, the user can see the information depicted in Figure 8 in the serial monitor.

### 5.2.2 2<sup>nd</sup> mode - Wifi access

Once the device has been connected to the EVISION access point, the user is redirected to the page depicted in Figure 9. In this page, the user is prompted to choose one of the available Wi-Fi networks in which the device is able to connect and connect to it by providing the network's Wi-Fi credentials. If the device is connected to another device providing access to a serial monitor, once the e-Vision device is connected to a Wi-Fi network, the serial monitor displays the information depicted in Figure 10.

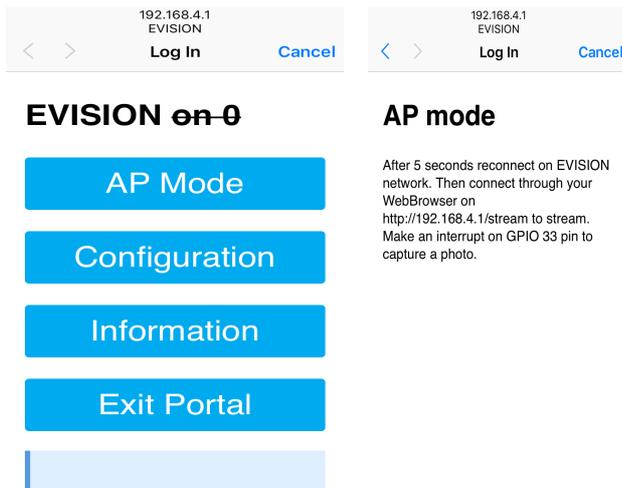


Fig. 7 Initialization of EVISION as an access point

```
rst:0x1 (POWERON_RESET),boot:0x13 (SPI_FAST_FLASH_BOOT)
configsip: 0, SPIWP:0xee
clk_drv:0x00,q_drv:0x00,d_drv:0x00,cs0_drv:0x00,hd_drv:0x00,wp_drv:0x00
mode:DIO, clock div:2
load:0x3fff0018,len:4
load:0x3fff001c,len:1216
ho 0 tail 12 room 4
load:0x40078000,len:9720
ho 0 tail 12 room 4
load:0x40080400,len:6364
entry 0x400806b8
Card initialised.
Camera Stream Ready!
Go to: http://192.168.4.1/stream to stream.
```

Fig. 8 Serial Screen after Connection to the EVISION access point

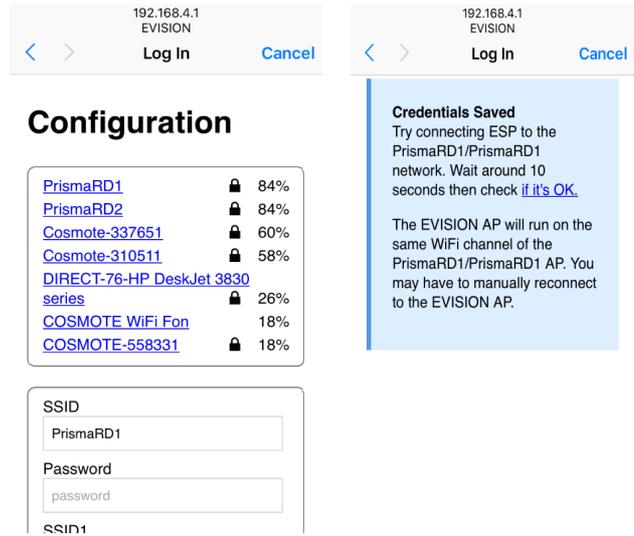


Fig. 9 EVISION connection to WiFi network

```
rst:0x1 (POWERON_RESET),boot:0x13 (SPI_FAST_FLASH_BOOT)
configsip: 0, SPIWP:0xee
clk_drv:0x00,q_drv:0x00,d_drv:0x00,cs0_drv:0x00,hd_drv:0x00,wp_drv:0x00
mode:DIO, clock div:2
load:0x3fff0018,len:4
load:0x3fff001c,len:1216
ho 0 tail 12 room 4
load:0x40078000,len:9720
ho 0 tail 12 room 4
load:0x40080400,len:6364
entry 0x400806b8
Card initialised.
Camera Stream Ready!
Go to: http://19.1.4.157/stream to stream.
```

Fig. 10 Serial Screen after connection to an existing Wi-Fi network

## 6 Conclusion

In this paper, two iterations of an integrated device employing a variety of sensors that can help visually impaired people is presented. One of the iterations employs COTS components, while the second one was an original device developed by the authors. Even though the second alternative is at its original state, both the hardware design and the development of integrated software allow integration as a solution for the needs of e-Vision. Completion tests and laboratory tests have showed that the solution can provide all the required functionalities. Furthermore, by adopting this alternative it is possible to significantly reduce the cost of the finished product. Of course, this solution would require some extra time to be integrated into a glass frame before it can be piloted. The examination of the adoption of this solution as a possible alternative to the sensor system in the glasses will be completed during the commercial evaluation of the system and the preparation of the corresponding design.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License.



## References

- Burton, M.J., Ramke, J., Marques, A.P., Bourne, R.R., Congdon, N., Jones, I., Tong, B.A.A., Arunga, S., Bachani, D., Bascaran, C. and Bastawrous, A., 2021. The Lancet Global Health Commission on Global Eye Health: vision beyond 2020. The Lancet Global Health.
- Bourne, R.R., 2021. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. The Lancet Global Health, 9(2), pp.e144-e160.
- www.transparencymarketresearch.com/assistive-technologies-visually-impaired-market.html
- J. P. Bigham, C. Jayant, A. Miller, B. White and T. Yeh, "VizWiz::LocateIt - enabling blind people to locate objects in their environment," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 2010.
- B. Schauerte, M. Martinez, A. Constantinescu and R. Stiefelhagen, "An Assistive Vision System for the Blind That Helps Find Lost

- Things," in *Computers Helping People with Special Needs. ICCHP 2012. Lecture Notes in Computer Science*, vol 7383, Springer, Berlin, Heidelberg, 2012, pp. 566-572.
6. A. Sirikham, W. Chiracharit and K. Chamnongthai, "Banknote and coin speaker device for blind people," in *International Conference on Advanced Communication Technology, ICACT*, Phoenix Park, South Korea, 2009.
  7. L. Chen, J. Su, M. Chen, W. Chang, C. Yang and C. Sie, "An Implementation of an Intelligent Assistance System for Visually Impaired/Blind People," in *2019 IEEE International Conference on Consumer Electronics (ICCE)*, 2019.
  8. R. Velazquez, E. E. Pissaloux, J. C. Guinot and F. Maingreud, "Walking Using Touch: Design and Preliminary Prototype of a Non-Invasive ETA for the Visually Impaired," in *Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE*, Shanghai, 2005.
  9. A. Suresh, D. Laha, D. Gaba and S. Bhambri, "Robot Intelligence Technology and Applications," in *Robot Intelligence Technology and Applications*, vol. 5, 2019, pp. 99-112.
  10. L. Dunai, B. D. Garcia, I. Lengua and G. Peris-Fajarnes, "3D CMOS sensor based acoustic object detection and navigation system for blind people," in *IECON Proceedings (Industrial Electronics Conference)*, Montreal, 2012.
  11. Tsantilas S., Spandonidis C., Giannopoulos F., Galiatsatos N., Karageorgiou D. and Giordamli C., "A Comparative Study of Wireless Communication Protocols in a Computer Vision System for improving the Autonomy of the Visually Impaired," *Journal of Engineering Science and Technology Review*, vol. 13, issue 1, pp. 72-76.
  12. Spandonidis C., Tsantilas S., Giannopoulos F., Giordamli C., Zyrichidou I. and Syropoulou P., "Design and Development of a New Cost-Effective Internet of Things Sensor Platform for Air Quality Measurements," *Journal of Engineering Science and Technology Review*, vol. 13, issue 6, pp. 81-91.
  13. Spandonidis C., Giordamli C., "Data-centric Operations in Oil & Gas Industry by the Use of 5G Mobile Networks and Industrial Internet of Things (IIoT)," in *Thirteenth International Conference on Digital Telecommunications (ICDT 2018)*, 2018, pp. 1-5, 2018.
  14. [14] Espressif Systems, *ESP32 Technical Reference Manual*. 202