Manuscript Details

Manuscript number	DISPLA_2017_8_R1
Title	Development and evaluation of a head-mounted display system based on stereoscopic images and depth algorithms for patients with visual impairment
Article type	Full length article

Abstract

In this work, we developed a wearable system using a commercial stereoscopic head- mounted display. We compared depth first-then contour (df-tc) and contour first-then depth (cf-td) algorithms in terms of images processed per second rate versus the window size and sweep values. Likewise, we performed a comparison of several edge detection methods in the same terms. The developed technical aid was clinically tested. We evaluated the preferred walking speed and the walking speed with and without the devices during three test circuits. For comparative analysis of anxiety levels, we recorded patients' heart rate data before, during and after the test. The system has proven its potential for enhancing the patients' mobility and reducing the level of anxiety related to movement activities.

Keywords	Head-mounted display, HMD, visual impairment, edge detection, multiplexing.
Manuscript category	Display Technologies

Corresponding Author	María Begoña Coco-Martín
Corresponding Author's Institution	Instituto de Oftalmobiología Aplicada (IOBA)
Order of Authors	María Begoña Coco-Martín, María Pichel-Mouzo, Juan Carlos Torres, Ricardo Vergaz, Rubén Cuadrado, Jose Pinto-Fraga, Rosa Coco
Suggested reviewers	David Piñero, Frank Eperjesi, Miguel J. Maldonado

Submission Files Included in this PDF

File Name [File Type]

Response to reviewers July 2018.docx [Response to Reviewers]

Manuscript-V7 with changes.docx [Highlights]

Manuscript-V7.final version.docx [Manuscript File]

Conflict of interest.docx [Conflict of Interest]

Valladolid, July 16th, 2018

To Prof. Richard H. Y. So

Editor in Chief

Displays

Dear Editor,

Thank you very much for your kind consideration about our manuscript. After resubmission of April, 2018, we answer here to a new Resubmission query of this month. We have found in it that the editors and reviewers insisted on some points that we had tried to correct in the last resubmission, but unfortunately, we failed to do it then. Besides, they added a sum of others. After a deep revision, we hope that we had achieved a manuscript written in a way that could be more suitable for Displays review.

Here you are the detailed answers to **the comments point-by-point from the editors and reviewers** of April's last revision:

Comments from the editors and reviewers:

-Organization of texts on the title page

To avoid errors during the typesetting process, some misplaced/unnecessary texts that follow the corresponding author info, such as Abbreviations and acronyms (there are some errors in the abbreviation list), Financial Support (This can be incorporated in the Acknowledgements section), Authors' Contributions (This can be given as a section to follow after the Acknowledgements section) should be deleted.

We have removed every unnecessary text, sending the financial support properly to the end of Acknowledgements section and Author's Contributions following it. The Abbreviations and acronyms appears as a section at the end of the manuscript, just for the editor decision in the case of finally inserting it.

-Language style:

For clarity, the following text revisions are suggested:

Abstract, line 1

"…..head mounted display."

For consistency throughout this paper, please change "…..head mounted display." to "…..head-mounted display." when "HMD" is not used. The same applies to other cases, if applicable.

Done

Abstract, line 7

“For comparative analysis of anxiety level……”

For clarity in the abstract’s context, the following minor revision is suggested:

“For comparative analysis of anxiety levels……”

Done

Introduction, paragraph 3, line 1

“The idea of using HMD for low-vision…..”

For clarity in the context, the following minor revision is suggested:

“The idea of using HMDs for low-vision…..”

Done

Introduction, paragraph 5, line 5

“….leading to an enhanced mobility”

For clarity in the context, the following minor revision is suggested:

“….leading to enhanced mobility of the patient.”

Done

Introduction, paragraph 5, line 6

“….this device alerted about the presence of possible collision objects using…”

For unambiguity in the context, the following minor revision is suggested:

“….this device alerted about the presence of possible collision-causing objects using…”

Done

Introduction, paragraph 6, line 3

“For such purpose,..”

For text fluency, the following minor revision is suggested:

“For such a purpose,..”

Done

Introduction, paragraph 6, line 6

“…..on mobility and anxiety level.”

For unambiguity in the context, the following minor revision is suggested:

“…..on mobility and anxiety levels.”

Done

Section 2, heading

"MATERIAL AND METHODS"

For clarity, the following minor revision is suggested:

"MATERIALS AND METHODS"

Done

Section 2, Introductory paragraph

For clarity, the following revised brief paragraph is suggested:

"This research followed the tenets of the Declaration of Helsinki. The research protocol of this study was approved by the local ethics committee and informed consent was obtained from each patient. The research was then conducted using the hardware systems and methods described below in this section."

Done

Section 3.2, paragraph 2

This paragraph is vague. Please further elaborate.

Done according following comments and adding this sentence at the end of the paragraph:

> *The images show that the best contour plots are achieved using this method, showing complete edges and low image noise.*

Section 3.2, paragraph 2, line 1

"…methods appears in Figure 7."

For text fluency, the following minor revision is suggested:

"…methods is shown in Figure 7."

Done

Section 3.2, paragraph 2, line 2

"Although the results seem very similar at first sight, they were very sensitive to image noise, being the Canny's method the most accurate of all of them."

For text fluency and consitency, the following minor revision is suggested:

"Although the results seem very similar at first sight, they are very sensitive to image noise, with the Canny method being the most accurate of all of them."

For consistency, "Canny method", not "Canny's method" is suggested.

Done

Section 3.2, paragraph 3, line 1

"….and their position,"

According to the context, "….and their positions," should be used.

Done

Section 3.4, paragraph 2, line 3

"Different edge detection algorithms, which were apparently very similar, were also proven, being the Canny algorithm an efficient way of outlining objects and suppressing noise, thereby easing the information processing by the user."

For clarity, please revise and elaborate the above sentence. It may need to be expressed with 2-3 sentences for easy comprehension.
Done. We changed the sentence by this one:

> *We have tested different detection algorithms. They were apparently very similar, but Canny algorithm proved the best one in terms of outlining objects and suppressing noise. This was the algorithm that the users preferred to get the best information from the projected image.*

Section 3.4, paragraph 3, line 3

"We performed a preliminary clinical validation of this system in 11 patients with retinal problems, a validation that confirmed the usefulness….."

For clarity, the following revision is suggested:

"With the participation of 11 patients with retinal problems, we performed a preliminary clinical validation. The validation confirmed the usefulness….."

Done

Section 3.4, paragraph 3, line 6

"….but with less level of change of HR…"

For clarity, the following revision is suggested:

"….but with lower degree in the change of HR…"

Done

Section 3.4, paragraph 3, line 8

"in patient's mobility with….."

For clarity, the following revision is suggested:

“in patients’ mobility with…..”

Done

Section 3.4, paragraph 3, line 8

“In our clinical study, the difference in the change in the HR when the patients performed the three circuits designed showed that the lower HR at the beginning of the test occurred when the patient used the HMD Star device. This suggests a potentially lower level of patient’s anxiety when performing different walking circuits using the HMD electronic device. It should be considered that worry in daily life might have substantial cardiac effects in addition to the effects of stressful events, especially in the form of work-related and anticipatory stress [14].”

The last sentence seems irrelevant to this study. It can be deleted.

This part of discussion should be further elaborated.

We agree with the reviewer, this sentence is not relevant for the rest of the study, and reference [14] has been deleted. Moreover, the information appearing in this paragraph seems redundant with the end of the previous one, so we have entirely removed it.

Conclusion

Please elaborate the conclusion section, both technically and literally.

Done. We have rewritten the conclusion section as follows:

We have designed and developed a new concept of HMD with stereoscopic vision using two cameras. The edges of the objects in the images are detected first using a Canny algorithm, the one selected by the users as the best in terms of quality of the processed image. The information of the distances to the objects is processed using a disparity algorithm applied only to the edges of the objects, a method that allows a real-time operation. The processed images show the edges of the objects with information of their distance inserted as color or thickness of the edge. The images are projected in the remaining visual field of the patient. We have shown in a clinical test that the HMD allows the patients’ mobility and obstacle detection ability, and reduces their anxiety level. Future research will focus on the improvement of the real-time operation, the portability and the comfort of the system, all of them items demanded by the patients at the end of the trial.

Please thoroughly proofread the whole manuscript and make appropriate revisions, including those not mentioned above, if any.

Done

-References

Ref. 2

Please check and confirm whether the citation info/source is correct.

It is. Check https://www.hopkinsmedicine.org/wilmer/about/employees/cvs/Massof_CV.pdf

For all that reasons we hope that all the changes performed can fit within your journal scope and fulfill all the proposed requirements and all your comments.

Looking forward to hearing from you soon,

Best regards,

Maria B Coco, PhD

Corresponding Author

Development and evaluation of a head-mounted display system based on stereoscopic images and depth algorithms for patients with visual impairment

María Begoña Coco-Martin;[1,3] María Pichel-Mouzo;[1] Juan Carlos Torres;[2] Ricardo Vergaz;[2] Rubén Cuadrado;[1] José Pinto-Fraga;[1] Rosa María Coco;[1]

[1]IOBA Instituto Universitario de Oftalmobiología Aplicada, Universidad de Valladolid, Paseo de Belén, 17, 47011 Valladolid, Spain.

[2]GDAF-UC3M, Grupo de Displays y Aplicaciones Fotónicas, Departamento de Tecnología Electrónica, Universidad Carlos III de Madrid, Av. de la Universidad, 30, 28911 Leganés, Spain.

[3] Research Unit, Hospital Clínico Universitario de Valladolid. Instituto de Ciencias de la Salud de Castilla y León. IECSCYL. Avenida Ramón y Cajal 7, 47011, Valladolid, Spain.

Corresponding author: María B Coco-Martín, PhD. Research Unit. Hospital Clínico Universitario de Valladolid. IECSCYL, Avda. Ramón y Cajal, 47011, Valladolid, Spain. Telephone: +34 983420000 . Email: mbcoco@uemc.es.

~~Abbreviations and acronyms: **ARMD** = age-related macular degeneration, **cf-td** = first-then depth, **df-tc** = depth first-then, **HMD** = head-mounted display, **HR** = heart rate, **PWS** = preferred walking speed, **VA** = visual acuity, **VF** = visual field, **WS** = walking speed.~~

~~Financial Support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. **There is no conflict of interest of any of the authors of this manuscript.**~~

Authors' contributions: Coco-Martín MB, Pichel-Mouzo M conceptualized and designed the study, coordinated and supervised data collection, carried out the initial analyses, drafted the initial manuscript, reviewed and revised the manuscript and approved the final manuscript as submitted. Torres JC, Vergaz R designed and developed the device, carried out the initial analyses, drafted the initial manuscript, reviewed and revised the manuscript, and approved the final manuscript as submitted. Cuadrado R, Coco RM and Pinto J coordinated and supervised data collection, critically reviewed the manuscript, and approved the final manuscript as submitted.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

ABSTRACT

In this work, we developed a wearable system using a commercial stereoscopic head-mounted display. We compared depth first-then contour (df-tc) and contour first-then depth (cf-td) algorithms in terms of images processed per second rate versus the window size and sweep values. Likewise, we performed a comparison of several edge detection methods in the same terms. The developed technical aid was clinically tested. We evaluated the preferred walking speed and the walking speed with and without the devices during three test circuits. For comparative analysis of anxiety levels, we recorded patients' heart rate data before, during and after the test. The system has proven its potential for enhancing the patients' mobility and reducing the level of anxiety related to movement activities.

Keywords: Head-mounted display, HMD, visual impairment, edge detection, multiplexing.

1. INTRODUCTION

Visual impairment is a global public health problem that affects 161 million people in the world. Visual impairment is distributed non-homogeneously across age groups, with the problem largely found in adults over 50 years of age and older [1]. Usually, most of the disabilities can be corrected with conventional eyeglasses or by medical intervention, but in severe impairments such as low-vision or moderate and severe visual disability, these solutions are not possible. The most common causes of moderate visual disability are glaucoma, macular degeneration, diabetic retinopathy, and trachoma.

There are a wide variety of low-vision aids in the market, both electro-optical and non-optical, which help patients to overcome some of the effects generated by their disability. In the last decade, the use of new technologies in assistive technologies has been increasing, with a growing market in the field of small display devices or micro-displays (less than 2 inches) for low-vision assistive devices, such as optical magnification systems implemented in the so-called Head--Mounted Displays (HMDs). These devices are placed close to the eye using a helmet or glasses as support, giving the impression to the users that they are in front of a big screen.

The idea of using HMDs for low-vision assistive devices was not initially well received mainly due to the potential adverse effect that this new technology might have on vision [2–4]. Almost a decade later, Peli conducted and published a research demonstrating that there was no harmful effect of the use of HMDs on vision [5]. Since the publication of this research, the use of new systems based on HMDs has become more popular to help visually impaired people, by using more efficiently the vision that they retain, magnifying the images or modifying their contrast.

Different experiences with HMDs in people with severe visual impairment have been previously reported [5]. Peli et al developed a device combining a see-through micro-display and a wide-angle camera in a concept named as multiplexing [7], designed to be a helpful tool for people suffering peripheral vision reduction or central field loss. These authors also described other devices in which the real image could be seen through the display [8], where the images captured by the camera were used to develop an edge detection algorithm. An overlapped dynamic image or white lines cartoon was presented over the real scene by the micro-display and, consequently, patient could see the portion of environment that was visually lost because of his/her disability [8]. Furthermore, this approach allowed patients to locate objects that were beyond their visual field [8]. Other wearable low-vision aids like that developed by Bryant et al were able to indicate the location of an imminent collision hazard [9].

All these HMD-based technical aids did not implement three-dimensional information when used a single camera, triggering an erroneous perception of objects and resulting in patient's localization mistakes. Balakrishnan et al developed a device allowing the patient to detect and distinguish the proximity of objects using two cameras and an HMD system, leading to enhanced mobility of the patient ~~leading to an enhanced mobility~~ [10]. Specifically, this device alerted about the presence of possible collision-causing objects ~~this device alerted about the presence of possible collision objects~~ using a stereo audio signal. Limna et al also developed an obstacle detection system based on a stereo vision device to be used with a stick [11].

In the current study, we have developed and clinically tested a wearable system based on a commercial augmented reality eyewear with a lightweight laptop located in a backpack. For such a purpose, we used a specific depth estimation algorithm and edge detection method. The purpose of the present work is to evaluate preliminarily in low-vision subjects the impact of a new augmented reality HMD system on mobility and anxiety levels. This new stereoscopic cost-effective system was developed to multiplex the real image with the one processed by edge enhancement algorithms, and using a parallax method after edge detection to determine distances and detect obstacles at a distance from 0.5 to 13 meters. Likewise, this system also provides information about the proximity of objects by color and thickness of the depicted contours overlapped in the real image.

2. MATERIALS AND METHODS

~~3.~~ This research followed the tenets of the Declaration of Helsinki. The research protocol of this study was approved by the local ethics committee and informed consent was obtained from each patient. The research was then conducted using the hardware systems and methods described below in this section~~An observational case series study was performed. Data gathering was carried out after the local ethics committee approved the study protocol and informed consent was obtained from each patient. This research followed the tenets of the Declaration of Helsinki~~.

~~3.1.~~2.1. *System design: definition of depth estimation algorithm*

Stereopsis is a characteristic of human vision that allows calculating the distance from the objects located around the system. As this is a stereoscopic vision-based low-vision optical aid, it uses a pair of cameras located in different positions from which pictures are taken. If a point in space is projected on both cameras in different positions, this difference is called "disparity", and allows retrieving the distance between the point and the camera system using triangulation analysis.

Prior to the execution of any of the following algorithms, it is mandatory to adjust the two field of views by a calibration of both camera positions. This could be performed using the two images that can be obtained pointing them to a chessboard-pattern test plate. Then, both hardware (position) and software corrections are made, the former by adjusting the position of the cameras, the latter by shifting some pixels in one image some positions towards a certain direction to match the equivalent ones in the other image.

~~3.1.1.~~2.1.1. *Depth estimation: depth first-then contour (df-tc)*

Let this algorithm be defined literally as ‘depth first-then contour’, or df-tc. One of the most common methods used for the estimation of the depth using a stereoscopic image is the calculation of correlation, which is based on the sum of absolute differences between both images in the pair, performing a search of the disparity between them, as expressed in equation (1):

$$SAD = \sum_{i=1}^{n}\sum_{j=1}^{m}\left|I_T(x+i,y+j) - I_I(x+i-d,y+j)\right| \forall d = 1,2,3,\ldots, sweep \qquad (1)$$

where $I_T(x,y)$, $I_I(x,y)$ are the values of the pixels located in the position (x,y) in the left and right images, respectively, and *nxm* is the size of a window where (x,y) is the central pixel. The disparity (d) between one pixel in the left picture and the one in the right picture is defined as *sweep,* the parameter to be minimized in the correlation.

With this algorithm, a dense depth map is obtained, as the one shown in Figure 1, where each color indicates a different depth value. However, typical errors in the calculation of distances cause the shapes of the objects on the depth map to have a different contour than the real ones when they are computed. As an example, in Figure 1 the crosswalk on the street is confused with the girl’s contour, giving a false impression of the final scene.

~~3.1.2.~~2.1.2. *Depth estimation: contour first-then depth (cf-td)*

The distance calculation process, which for humans is almost trivial due to the speed and efficiency of the brain, is not so trivial for a computer due to the requirements of very high computational work and consumption of resources. The calculations of the disparity require the use of a high computing power. In many applications, an approximate distance estimated for an object may be sufficient, which can be computed only in a representative area of it, thus reducing the computing time. Therefore, it is necessary to clearly determine the feature of each element that contains the necessary information about its distance. The edge of the object is the most relevant shape feature that can be used both to retrieve the distance and to maintain the information about it at the same time. We developed another algorithm where the depth estimation was only computed in the objects edges, and therefore a reduced version of the disparity map was computed, involving only the contour pixels. Let this algorithm to be named 'contour first-then depth'or cf-td.

Figure 2 demonstrates how images processed in this way provide a clearer result than using only correlation disparity as in Figure 1. For a better visualization of the association of the contours to objects, the real scene is overlapped in the image.

df-tc and cf-tc algorithms were compared in terms of images processed per second rate versus the *window size* and with *sweep* values of 7, 10, 13 and 25. This rate is the inverse magnitude of the processing time, so the higher the rate the better for the implementation in real time systems.

~~3.1.3.~~2.1.3. *Definition of the edge detection method*

We implemented several edge detection methods (Canny, Prewitt, Sobel and Roberts) and compared them in terms of computing time for various *sweep* and *window size* values. As Maini and Aggarwal demonstrated that Prewitt, Sobel and Roberts methods were inaccurate and had sensitivity to noise and that the Canny~~'s~~ method was time consuming, a comparison among these methods were done prior to the definition of the developed system [6].

~~3.2.~~2.2. ***Configuration and specification of the vision aid system***

The technical aid developed in this work as a proof of concept for the implementation of the selected depth estimation algorithm and edge estimation method was based on a small, lightweight, augmented reality eyewear that was developed and manufactured by Vuzix. The system was powered by standard AC current or optionally by two replaceable/rechargeable batteries, allowing autonomy for at least 4 hours. To capture and process images, a handy laptop of 1.34 kg, with Intel Core i5 processor, was used. In addition, the laptop was equipped with at least 2 USBs and VGA video output, all of them necessary to connect to the Vuzix eyewear.

The HMD model was a Wrap 920AR with two high-resolution 640 x 480 LCD displays and two discrete VGA video cameras, providing images from 160 x 120 pixels to 640 x 480 pixels, but low resolution is enough for detection and distance information purposes. The system allowed the patient to manage in an indoor environment, pressing a button, entering in the elevator, searching his/her way all through a corridor (Figure 3), or detecting some obstacles and the basic lines of a street in outdoor conditions. The Wrap model provides a virtual image of apparent real size. Likewise, we also used a STAR model in the clinical testing, which provides a virtual image overlapped on a smaller real image.

~~3.3.~~2.3. ***Clinical test of the vision aid system***

~~3.3.1.~~2.3.1. *Patients*

We performed a clinical testing of the system developed using the Wrap and STAR models with the participation of visually impaired people due to retinopathy and other retinal diseases.

11 patients, 8 with retinitis pigmentosa and 3 with age-related macular degeneration (ARMD), were enrolled in this clinical testing. Inclusion criteria were visual field (VF) of 20° or less and best spectacle corrected visual acuity (VA) of 20/50 or less in the eye with the best vision. Patients with physical/mental problems or physical disabilities were excluded. All procedures were performed according to the Declaration of Helsinki. The study was approved by the local ethics committee and each patient signed an informed consent.

~~3.3.2.~~2.3.2. *Evaluation*

The preferred walking speed (PWS) of each subject was determined without devices prior to the measurement in the circuits, and afterwards with the respective electronic prototypes Star 1200 and Wrap 920AR, by measuring the time in seconds needed to walk straight a distance of 20 meters on a smooth pavement without obstacles. We performed this test three times in each patient with and without devices, being the mean of the three PWS measurements considered for the analysis.

We designed three 40-meter circuits of medium difficulty, each one different for each device tested (Wrap and STAR) to avoid a potential learning effect. We measured the time needed by each patient to finish each circuit with obstacles, adding a certain penalization time for each error done during the test, according to the criteria in a penalties table (computed as a compensation for the theoretical gain of time associated to each error done by the patient). This real time for finishing the circuit was used to estimate the walking speed (WS) in each circuit with and without the HMDs. Heart rate (HR) was also recorded before, during and after the test with each device. We compared WS with the PWS of each patient.

We optimized the parameters for the Wrap and STAR models and kept them constant throughout the study: 800 x 600 pixels resolution, equal thickness contours, 80 x 105 image size, black and white, contour by Canny algorithm (0.4), contours color value for each distance (red 78.1 cm, 96.9 cm yellow green, 131.3 cm, blue 220.7 cm), window size 3x3 and scan size 13.

~~3.3.3.~~2.3.3. *Statistical analysis*

We collected all data recorded during the measurements into a database and statistical calculations were performed using commercial statistical software (SPSS version 18.0 for Windows). We calculated the mean and standard deviation for a standard data distribution, while the median and interquartile ranges were calculated in case of non-parametric distributions. For all statistical tests, 2-tailed $p<0.05$ was considered as significant. To compare the results of quantitative variables we used the Wilcoxon test.

4~~.~~3. Results and Discussion

~~4.1.~~3.1. Depth first-then contour vs contour first-then depth

The comparative analysis of df-tc and cf-tc algorithms in terms of images processed per second rate versus the window size and with sweep values of 7, 10, 13 and 25 (Figure 4) revealed the following:

- In all cases, df-tc algorithm is 3 to 11 times slower than the cf-td depending on the resolution, sweep and window size.
- There is a linear decrease of the processed rate of images with the size of the window applied, higher for higher sweep parameters, as expected. Only the sweep 7 algorithm overcomes the 25 images per second required for an actual real time application video rate.
- Any real-time application should be at least working at 16 frames per second, a rate achieved in all the cases by the cf-td algorithm at 80 x 105 pixels resolution.
- By increasing the resolution to 144 x 192 pixels, the processing times increased even over the double than for 80 x 105 pixels.

Analysing the results of the graphs shown above the conclusion is that the smaller the sweep value and window size, the smaller the computing time. However, decreasing the sweep value or window size does not lead to worse results in the depth calculation (Figure 5).

~~4.2.~~3.2. Comparative analysis of edge detection methods

A linear decrease of the frames processed per second with the window size was obtained (Figure 6), with higher slope for higher sweep parameters. This decrease seems to depend mainly on the sweep rather than on the algorithm used for edge detection. The comparative analysis among methods revealed that the Canny~~'s~~ algorithm was the slowest method, but with a rate over 16 images per second for 80 x 105 pixel resolution in all cases. The differences among algorithms decreased when the window size parameter increased.

A comparison among Prewitt, Sobel, Roberts, and Canny methods ~~appears~~ is shown in Figure 7. Although the results seem very similar at first sight, they are very sensitive to image noise, with the Canny method being the most accurate of all of them~~Although the results seem very similar at first sight, they were very sensitive to image noise, being the Canny's method the most accurate of all of them~~. The images show that the best contour plots are achieved using this method, showing complete edges and low image noise.

Finally, in order to improve the ability of perceiving objects and their positions, and to minimize the influence of color detection skills of the patient, we performed an algorithm to project the information relating to depth/distance by a variation of thickness and color of the edges, as shown in Figure 8.

~~4.3.~~3.3. ***Results of the clinical study***

The percentage of women included in the sample was 54.55% (n=6), compared to 45.45% of males (n=5). The mean age of subjects was 52.0 ± 23.3 years. The average of the PWS without any device was 1.33 ± 0.28 m/s, 1.14 ± 0.27 m/s with the Star prototype, and 0.96 ± 0.3 m/s with the Wrap device (Figure 9). PWS without device was significantly higher than the value obtained with Wrap (p=0.001) and STAR (p=0.0174) electronic devices. The WS in the circuit without device was 0.76 ± 0.18 m/s, 0.57± 0.18 m/s with the Star prototype and 0.44± 0.12 m/s with the Wrap device. Comparing the measures in the three conditions, the WS without device was significantly higher than that observed with the STAR (p=0.0498) and Wrap electronic prototypes (p=0.002) (Figure 9). Mean differences among PWS and WS were 0.57 ± 0.22 m/s without device, 0.57± 0.20 m/s with the Star prototype and 0.51 ± 0.20 m/s with the Wrap device. The difference among PWS and WS with and without devices was statistically significant (p<0.001).

Without devices, the HR is significantly higher during testing compared to baseline (78.82 ± 15.52 vs. 70.64 ± 10.35 lpm, p=0.006) or the end of the circuit (74.82 ± 11.63 vs. 70.64 ± 10.35 lpm, p=0.004). With the Star prototype, the HR was significantly higher during the test compared to baseline (76.36 ± 16.30 vs. 69.82 ± 16.41 lpm, p=0.01). With Wrap device, the HR was also significantly higher during the test compared to baseline (80.00 ± 14.25 vs. 70.09 ± 9.59 lpm, p=0.01). Comparing the measurements in the 3 conditions, the HR at the beginning of the test was significantly lower with Star device compared to the non-wearing device condition (p=0.01). Mean increase in the HR during the test circuit was 9.45 ± 7.63 lpm without device, 6.55 ± 10.54 lpm with the Star prototype and 4.18 ± 10.15 lpm with the Wrap device.

~~4.4.~~3.4. Discussion

According to our calculations, the correlation based on the sum of absolute differences between stereoscopic images was the algorithm for the calculation of depth with the lowest computing cost, and for this reason, we selected it for the system developed in the current study. However, any error in the calculation of distances was found to change the shape of the depth map objects regarding real objects when a dense disparity map was used. To solve this problem, we implemented an algorithm to compute the distance from the objects using only their edges, performing the edge detection prior to the disparity computations, a method that also improved the computing times.

Our simulations show that the calculation times were mainly impacted by factors such as the image size, the sweep value and the used window size, but the edge detection method used should be also considered. We have tested different detection algorithms. They were apparently very similar, but Canny algorithm proved the best one in terms of~~Different~~

edge detection algorithms, which were apparently very similar, were also proven, being the Canny algorithm an efficient way of outlining objects and suppressing noise. This was the algorithm that the users preferred to get the best information from the projected image., thereby easing the information processing by the user.

Considering an optimized real-time depth estimation algorithm according to our simulations and the Canny edge detection method, a proof of concept system was developed using a commercial HMD and a laptop. With the participation of 11 patients with retinal problems, we performed a preliminary clinical validation. The validation confirmed the usefulness We performed a preliminary clinical validation of this system in 11 patients with retinal problems, a validation that confirmed the usefulness of the system to maintain acceptable levels of PWS and WS during three test circuits, but with lower degree in the change of HR but with less level of change of HR associated to the performance of these walking tasks. This is consistent with the results of other studies reporting improvements in patient's' mobility with other HMDs [12,13]. In our clinical study, the difference in the change in the HR when the patients performed the three circuits designed showed that the lower HR at the beginning of the test occurred when the patient used the HMD Star device. This suggests a potentially lower level of patient's anxiety when performing different walking circuits using the HMD electronic device. It should be considered that worry in daily life might have substantial cardiac effects in addition to the effects of stressful events, especially in the form of work-related and anticipatory stress [14].

5.4. Conclusion

We have designed and developed a new concept of HMD with stereoscopic vision using two cameras. The edges of the objects in the images are detected first using a Canny algorithm, the one selected by the users as the best in terms of quality of the processed image. The information of the distances to the objects is processed using a disparity algorithm applied only to the edges of the objects, a method that allows a real-time operation. The processed images show the edges of the objects with information of their distance inserted as color or thickness of the edge. The images are projected in the remaining visual field of the patient. We have shown in a clinical test that the HMD allows the patients' mobility and obstacle detection ability, and reduces their anxiety level.Summing up, a new concept of HMD with a three-dimensional camera integrated has been developed that shows not only the edges of the obstacles and objects, but also information about their proximity by the depicted colors or edge thickness in the finally projected images. We have shown the potential of this HMD has shown to improve the patient's mobility and to reduce the level of anxiety related to movement activities. Our next Future research will focus on the improvement of the real-time operation, the portability and the comfort of the system, all of them items demanded by the patients at the end of the trial.

Acknowledgements

We thank the Comunidad Autónoma de Madrid and University Carlos III de Madrid for funding in the Project ATIDiVisTA (Ref. CCG10-UC3M/TIC-4787), and the Comunidad Autónoma de Madrid for the Project FACTOTEM-2 (grant. nº S2009/ESP-1781) and SINFOTON (grant nº S2013/MIT-2790).

There is no conflict of interest of any of the authors of this manuscript.

Authors' contributions: Coco-Martín MB, Pichel-Mouzo M conceptualized and designed the study, coordinated and supervised data collection, carried out the initial analyses, drafted the initial manuscript, reviewed and revised the manuscript and approved the final manuscript as submitted. Torres JC, Vergaz R designed and developed the device, carried out the initial analyses, drafted the initial manuscript, reviewed and revised the manuscript, and approved the final manuscript as submitted. Cuadrado R, Coco RM and Pinto J coordinated and supervised data collection, critically reviewed the manuscript, and approved the final manuscript as submitted.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

~~Conflicts of Interest: Coco-Martín MB, None; Pichel-Mouzo M, None; Torres JC, None; Vergaz R, None; Cuadrado R, None; Pinto-Fraga J, None; Coco RM, None.~~

References

[1] S. Resnikoff, D. Pascolini, D. Etya'ale, I. Kocur, R. Pararajasegaram, G.P. Pokharel, S.P. Mariotti, Global data on visual impairment in the year 2002, Bull. World Health Organ. 82 (2004) 844–51.

[2] R. W. Massof, Electro-optical head-mounted low-vision enhancement, Pract. Optom. 9 (6), (1998) 214–220.

[3] J.P. Wann, S. Rushton, M. Mon-Williams, Natural problems for stereoscopic depth perception in virtual environments, Vision Res. 35 (1995) 2731–2736.

[4] M. Mon-Williams, J.P. Warm, S. Rushton, Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display, Ophthalmic Physiol. Opt. 13 (1993) 387–391.

[5] E. Peli, The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display, Vision Res. 38 (1998) 2053–2066.

[6] R. Maini, H. Aggarwal, Study and comparison of various image edge detection techniques, Int. J. Image Process. 3 (2010) 1–12.

[7] E. Peli, G. Luo, A. Bowers, N. Rensing, Development and evaluation of vision multiplexing devices for vision impairments, Int. J. Artif. Intell. Tools, 18 (2009) 365–378.

[8] E. Peli, G. Luo, A. Bowers, N. Rensing, Applications of augmented vision head-mounted systems in vision rehabilitation, J. Soc. Inf. Disp. 15 (2007) 1037–1045.

[9] R.C. Bryant, C.M. Lee, R.A. Burstein, E.J. Seibel, Engineering a low-cost wearable low vision aid based on retinal light scanning, SID Symp. Dig. Tech. Pap. 35 (2004) 1540–1543.

[10] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, S. Yaacob, Wearable real-time stereo vision for the visually impaired, Eng. Lett. 14 (2007) 6–14.

[11] T. Limna, P. Tandayya, N. Suvanvorn, Low-cost stereo vision system for supporting the visually impaired's walk, in: 3rd Int. Conv. Rehabil. Eng. Assist. Technol., 2009. Singapore.

[12] M.D. Peláez-Coca, F. Vargas-Martín, S. Mota, J. Díaz, E. Ros-Vidal, A versatile optoelectronic aid for low-vision patients, Ophthalmic Physiol. Opt. 29 (2009) 565–572.

[13] S.L. Hicks, I. Wilson, L. Muhammed, J. Worsfold, S.M. Downes, C. Kennard, A depth-based head-mounted visual display to aid navigation in partially sighted individuals, PLoS One. 8 (2013) e67695.

[14] S. Pieper, J.F. Brosschot, R. van der Leeden, J.F. Thayer, Cardiac effects of momentary assessed worry episodes and stressful events, Psychosom. Med. 69 (2007) 901–909.

Figures

Figure 1. Depth estimation: depth first-then contour analysis. Left: Original picture. Right: final result from adding depth information (via color assignation to the objects, with a scale on the right side) to the disparity map. Image resolution 144 x 192 pixels, sweep 27, and window size 9 x 9.

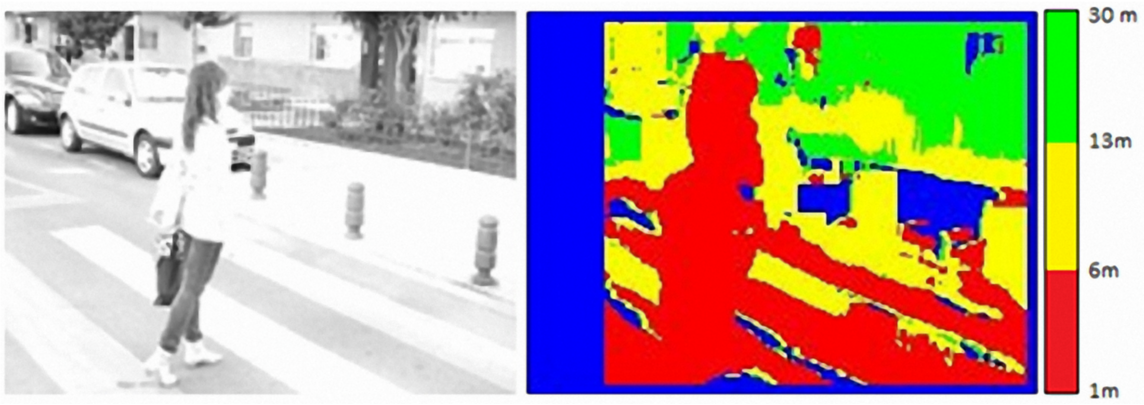


Figure 2. Depth estimation: contour first-then depth analysis. Left: depth map. Right: final result of overlapping depth information to the image. Image resolution 144 x 192 pixels, sweep 27, and window size 9 x 9.

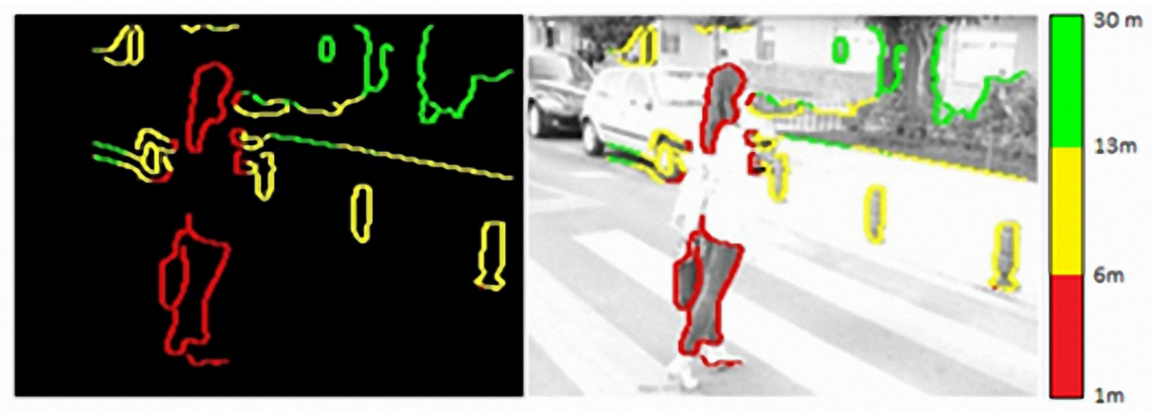
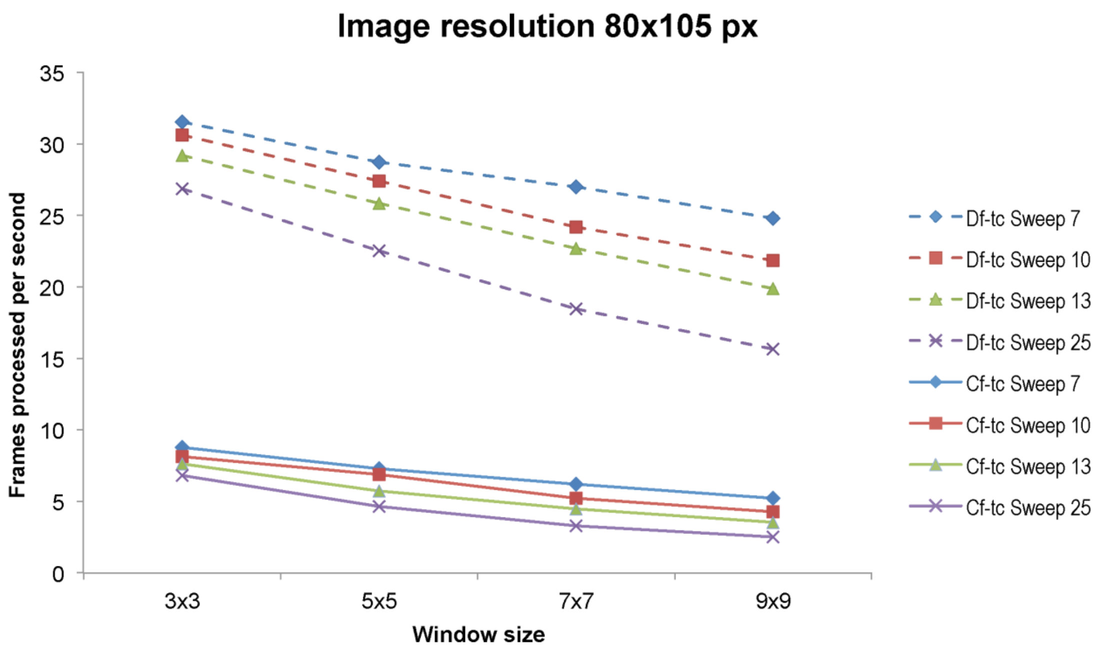


Figure 3. A non-visually disabled person wearing the prototype system. The prototype items are shown in the bottom of figure.

Figure 4. Computing time between traditional and new method for images with different resolution: upper plot, 80 x 105 pixels, and 144 x 192 pixels for the plot below.



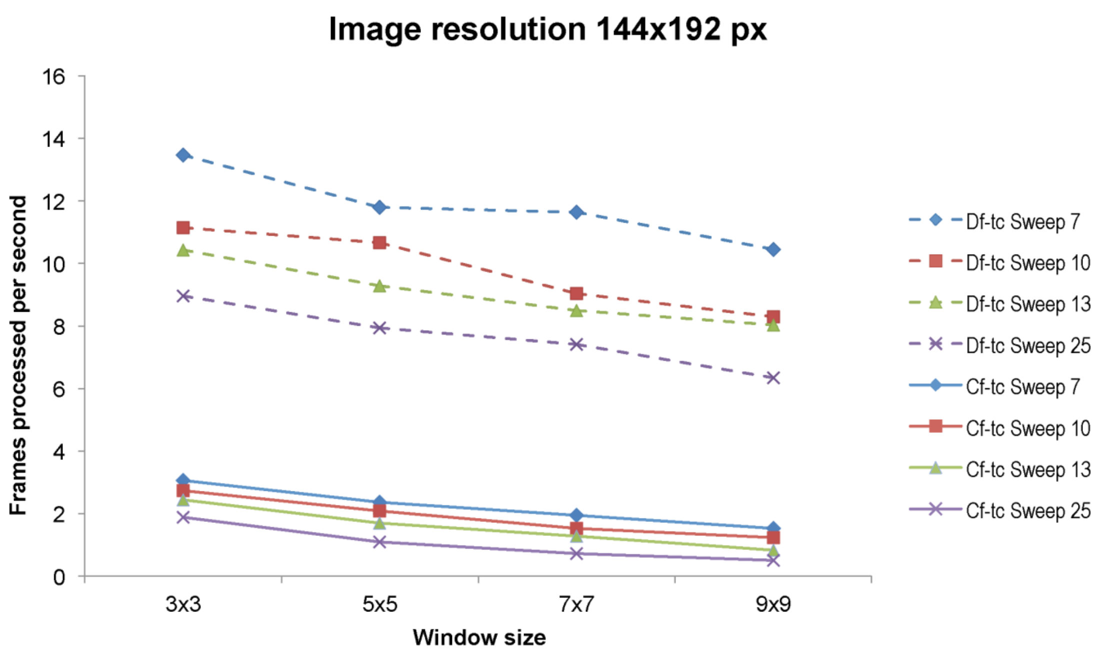


Figure 5. An example of a woman crossing a crosswalk using different calculation parameters: (a) Window size 3 x 3 and Sweep 7, (b) Window size 3 x 3 and Sweep 27, (c) Window size 9 x 9 and Sweep 7, (d) Window size 9 x 9 and Sweep 27 for image resolution of 120 x 160 pixels.

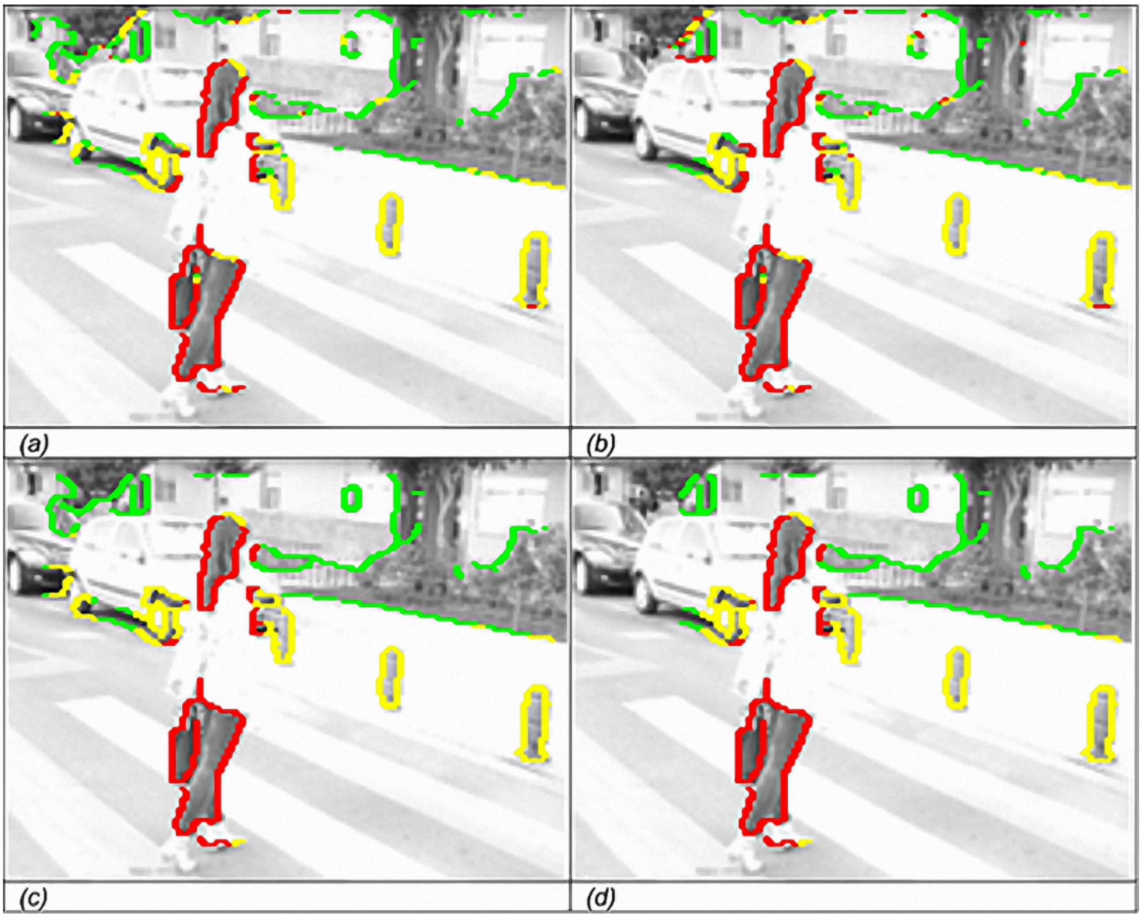


Figure 6. Comparison calculation times by using different detection methods for imaging contours. 80 x 105 pixels in the image, blue means a parameter sweep value of 7, red 13 and green 25.

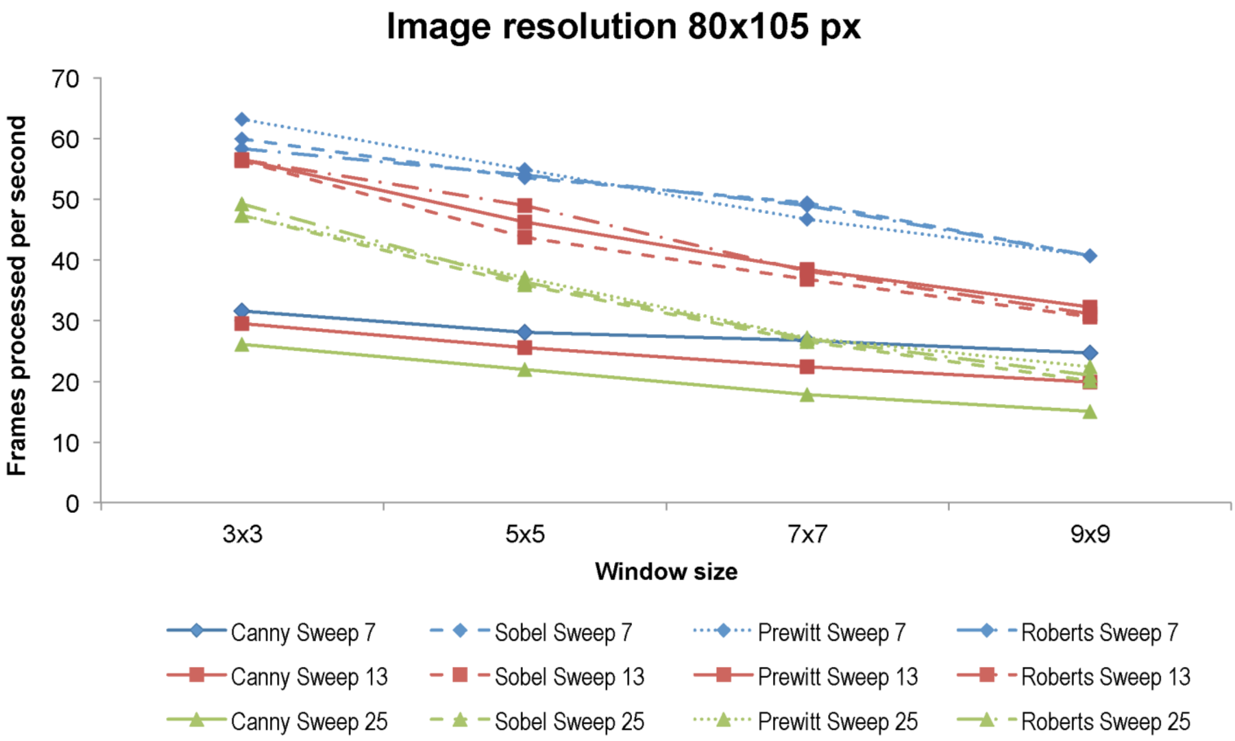


Figure 7. Comparison among different edge detection algorithms: (a) Canny Method, (b) Sobel Method, (c) Prewitt Method and (d) Roberts Method for image resolution of 80 x 105 pixels, sweep 13 and window size 5 x 5.

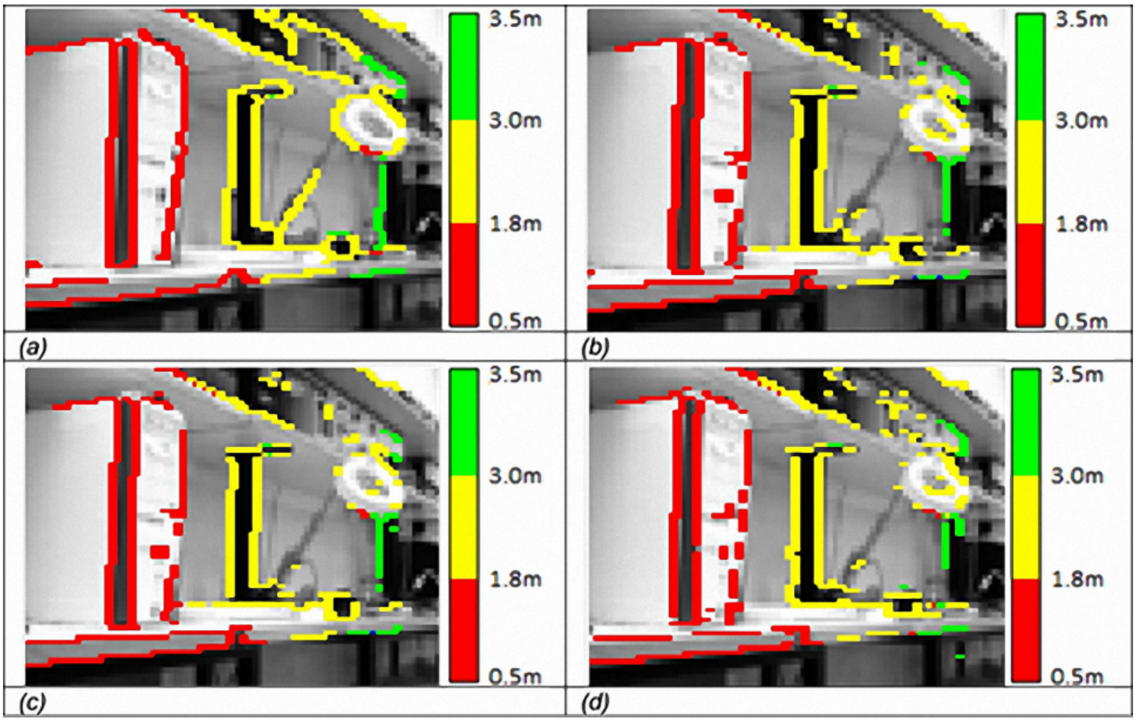


Figure 8. Variation of thickness and color in the edges.

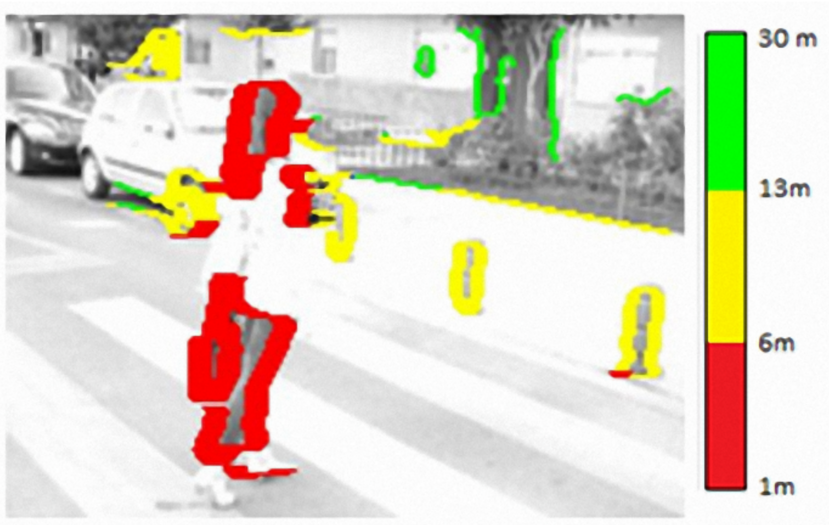
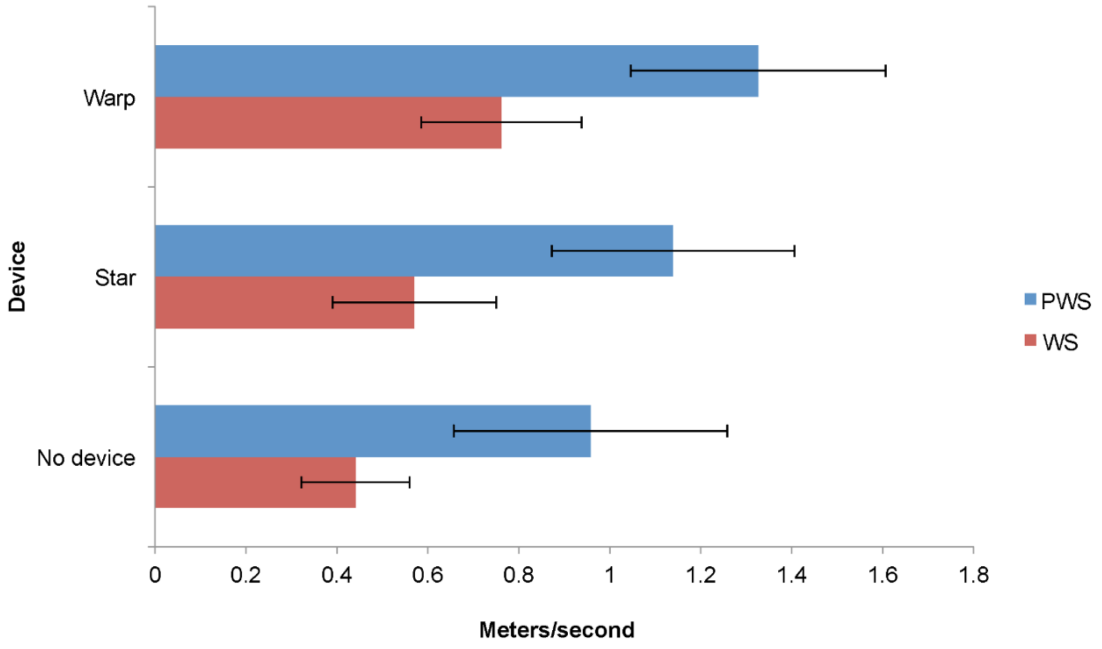


Figure 9. Preferred walking speed (PWS) and walking speed (WS) in the circuits without devices and with the electronic prototypes Star 1200 and Wrap 920AR.



Abbreviations and acronyms: **ARMD** = age-related macular degeneration, **cf-td** = contour first-then depth, **df-tc** = depth first-then contour, **HMD** = head-mounted display, **HR** = heart rate, **PWS** = preferred walking speed, **VA** = visual acuity, **VF** = visual field, **WS** = walking speed.

Development and evaluation of a head-mounted display system based on stereoscopic images and depth algorithms for patients with visual impairment

María Begoña Coco-Martin;[1,3] *María Pichel-Mouzo;*[1] *Juan Carlos Torres;*[2] *Ricardo Vergaz;*[2] *Rubén Cuadrado;*[1] *José Pinto-Fraga;*[1] *Rosa María Coco;*[1]

[1]IOBA Instituto Universitario de Oftalmobiología Aplicada, Universidad de Valladolid, Paseo de Belén, 17, 47011 Valladolid, Spain.

[2]GDAF-UC3M, Grupo de Displays y Aplicaciones Fotónicas, Departamento de Tecnología Electrónica, Universidad Carlos III de Madrid, Av. de la Universidad, 30, 28911 Leganés, Spain.

[3] Research Unit, Hospital Clínico Universitario de Valladolid. Instituto de Ciencias de la Salud de Castilla y León. IECSCYL. Avenida Ramón y Cajal 7, 47011, Valladolid, Spain.

Corresponding author: María B Coco-Martín, PhD. Research Unit. Hospital Clínico Universitario de Valladolid. IECSCYL, Avda. Ramón y Cajal, 47011, Valladolid, Spain. Telephone: +34 983420000 . Email: mbcoco@uemc.es.

ABSTRACT

In this work, we developed a wearable system using a commercial stereoscopic head-mounted display. We compared depth first-then contour (df-tc) and contour first-then depth (cf-td) algorithms in terms of images processed per second rate versus the window size and sweep values. Likewise, we performed a comparison of several edge detection methods in the same terms. The developed technical aid was clinically tested. We evaluated the preferred walking speed and the walking speed with and without the devices during three test circuits. For comparative analysis of anxiety levels, we recorded patients' heart rate data before, during and after the test. The system has proven its potential for enhancing the patients' mobility and reducing the level of anxiety related to movement activities.

Keywords: Head-mounted display, HMD, visual impairment, edge detection, multiplexing.

1. INTRODUCTION

Visual impairment is a global public health problem that affects 161 million people in the world. Visual impairment is distributed non-homogeneously across age groups, with the problem largely found in adults over 50 years of age and older [1]. Usually, most of the disabilities can be corrected with conventional eyeglasses or by medical intervention, but in severe impairments such as low-vision or moderate and severe visual disability, these solutions are not possible. The most common causes of moderate visual disability are glaucoma, macular degeneration, diabetic retinopathy, and trachoma.

There are a wide variety of low-vision aids in the market, both electro-optical and non-optical, which help patients to overcome some of the effects generated by their disability. In the last decade, the use of new technologies in assistive technologies has been increasing, with a growing market in the field of small display devices or micro-displays (less than 2 inches) for low-vision assistive devices, such as optical magnification systems implemented in the so-called Head-Mounted Displays (HMDs). These devices are placed close to the eye using a helmet or glasses as support, giving the impression to the users that they are in front of a big screen.

The idea of using HMDs for low-vision assistive devices was not initially well received mainly due to the potential adverse effect that this new technology might have on vision [2–4]. Almost a decade later, Peli conducted and published a research demonstrating that there was no harmful effect of the use of HMDs on vision [5]. Since the publication of this research, the use of new systems based on HMDs has become more popular to help visually impaired people, by using more efficiently the vision that they retain, magnifying the images or modifying their contrast.

Different experiences with HMDs in people with severe visual impairment have been previously reported [5]. Peli et al developed a device combining a see-through micro-display and a wide-angle camera in a concept named as multiplexing [7], designed to be a helpful tool for people suffering peripheral vision reduction or central field loss. These authors also described other devices in which the real image could be seen through the display [8], where the images captured by the camera were used to develop an edge detection algorithm. An overlapped dynamic image or white lines cartoon was presented over the real scene by the micro-display and, consequently, patient could see the portion of environment that was visually lost because of his/her disability [8]. Furthermore, this approach allowed patients to locate objects that were beyond their visual field [8]. Other wearable low-vision aids like that developed by Bryant et al were able to indicate the location of an imminent collision hazard [9].

All these HMD-based technical aids did not implement three-dimensional information when used a single camera, triggering an erroneous perception of objects and resulting in patient's localization mistakes. Balakrishnan et al developed a device allowing the patient to detect and distinguish the proximity of objects using two cameras and an HMD system, leading to enhanced mobility of the patient [10]. Specifically, this device alerted about the presence of possible collision-causing objects using a stereo audio signal. Limna et al also developed an obstacle detection system based on a stereo vision device to be used with a stick [11].

In the current study, we have developed and clinically tested a wearable system based on a commercial augmented reality eyewear with a lightweight laptop located in a backpack. For such a purpose, we used a specific depth estimation algorithm and edge detection method. The purpose of the present work is to evaluate preliminarily in low-vision subjects the impact of a new augmented reality HMD system on mobility and anxiety levels. This new stereoscopic cost-effective system was developed to multiplex the real image with the one processed by edge enhancement algorithms, and using a parallax method after edge detection to determine distances and detect obstacles at a distance from 0.5 to 13 meters. Likewise, this system also provides information about the proximity of objects by color and thickness of the depicted contours overlapped in the real image.

2. MATERIALS AND METHODS

This research followed the tenets of the Declaration of Helsinki. The research protocol of this study was approved by the local ethics committee and informed consent was obtained from each patient. The research was then conducted using the hardware systems and methods described below in this section.

2.1. System design: definition of depth estimation algorithm

Stereopsis is a characteristic of human vision that allows calculating the distance from the objects located around the system. As this is a stereoscopic vision-based low-vision optical aid, it uses a pair of cameras located in different positions from which pictures are taken. If a point in space is projected on both cameras in different positions, this difference is called "disparity", and allows retrieving the distance between the point and the camera system using triangulation analysis.

Prior to the execution of any of the following algorithms, it is mandatory to adjust the two field of views by a calibration of both camera positions. This could be performed using the two images that can be obtained pointing them to a chessboard-pattern test plate. Then, both hardware (position) and software corrections are made, the former by adjusting the position of the cameras, the latter by shifting some pixels in one image some positions towards a certain direction to match the equivalent ones in the other image.

2.1.1. Depth estimation: depth first-then contour (df-tc)

Let this algorithm be defined literally as 'depth first-then contour', or df-tc. One of the most common methods used for the estimation of the depth using a stereoscopic image is the calculation of correlation, which is based on the sum of absolute differences between both images in the pair, performing a search of the disparity between them, as expressed in equation (1):

$$SAD = \sum_{i=1}^{n}\sum_{j=1}^{m}\left|I_T(x+i,y+j) - I_I(x+i-d,y+j)\right| \forall d = 1,2,3,\ldots, sweep \qquad (1)$$

where $I_T(x,y)$, $I_I(x,y)$ are the values of the pixels located in the position (x,y) in the left and right images, respectively, and *nxm* is the size of a window where (x,y) is the central pixel. The disparity (d) between one pixel in the left picture and the one in the right picture is defined as *sweep,* the parameter to be minimized in the correlation.

With this algorithm, a dense depth map is obtained, as the one shown in Figure 1, where each color indicates a different depth value. However, typical errors in the calculation of distances cause the shapes of the objects on the depth map to have a different contour than the real ones when they are computed. As an example, in Figure 1 the crosswalk on the street is confused with the girl's contour, giving a false impression of the final scene.

2.1.2. Depth estimation: contour first-then depth (cf-td)

The distance calculation process, which for humans is almost trivial due to the speed and efficiency of the brain, is not so trivial for a computer due to the requirements of very high computational work and consumption of resources. The calculations of the disparity require the use of a high computing power. In many applications, an approximate distance estimated for an object may be sufficient, which can be computed only in a representative area of it, thus reducing the computing time. Therefore, it is necessary to clearly determine the feature of each element that contains the necessary information about its distance. The edge of the object is the most relevant shape feature that can be used both to retrieve the distance and to maintain the information about it at the same time. We developed another algorithm where the depth estimation was only computed in the objects edges, and therefore a reduced version of the disparity map was computed, involving only the contour pixels. Let this algorithm to be named 'contour first-then depth'or cf-td.

Figure 2 demonstrates how images processed in this way provide a clearer result than using only correlation disparity as in Figure 1. For a better visualization of the association of the contours to objects, the real scene is overlapped in the image.

df-tc and cf-tc algorithms were compared in terms of images processed per second rate versus the *window size* and with *sweep* values of 7, 10, 13 and 25. This rate is the inverse magnitude of the processing time, so the higher the rate the better for the implementation in real time systems.

2.1.3. Definition of the edge detection method

We implemented several edge detection methods (Canny, Prewitt, Sobel and Roberts) and compared them in terms of computing time for various *sweep* and *window size* values. As Maini and Aggarwal demonstrated that Prewitt, Sobel and Roberts methods were inaccurate and had sensitivity to noise and that the Canny method was time consuming, a comparison among these methods were done prior to the definition of the developed system [6].

2.2. ***Configuration and specification of the vision aid system***

The technical aid developed in this work as a proof of concept for the implementation of the selected depth estimation algorithm and edge estimation method was based on a small, lightweight, augmented reality eyewear that was developed and manufactured by Vuzix. The system was powered by standard AC current or optionally by two replaceable/rechargeable batteries, allowing autonomy for at least 4 hours. To capture and process images, a handy laptop of 1.34 kg, with Intel Core i5 processor, was used. In addition, the laptop was equipped with at least 2 USBs and VGA video output, all of them necessary to connect to the Vuzix eyewear.

The HMD model was a Wrap 920AR with two high-resolution 640 x 480 LCD displays and two discrete VGA video cameras, providing images from 160 x 120 pixels to 640 x

480 pixels, but low resolution is enough for detection and distance information purposes. The system allowed the patient to manage in an indoor environment, pressing a button, entering in the elevator, searching his/her way all through a corridor (Figure 3), or detecting some obstacles and the basic lines of a street in outdoor conditions. The Wrap model provides a virtual image of apparent real size. Likewise, we also used a STAR model in the clinical testing, which provides a virtual image overlapped on a smaller real image.

2.3. Clinical test of the vision aid system

2.3.1. Patients

We performed a clinical testing of the system developed using the Wrap and STAR models with the participation of visually impaired people due to retinopathy and other retinal diseases.

11 patients, 8 with retinitis pigmentosa and 3 with age-related macular degeneration (ARMD), were enrolled in this clinical testing. Inclusion criteria were visual field (VF) of 20° or less and best spectacle corrected visual acuity (VA) of 20/50 or less in the eye with the best vision. Patients with physical/mental problems or physical disabilities were excluded. All procedures were performed according to the Declaration of Helsinki. The study was approved by the local ethics committee and each patient signed an informed consent.

2.3.2. Evaluation

The preferred walking speed (PWS) of each subject was determined without devices prior to the measurement in the circuits, and afterwards with the respective electronic prototypes Star 1200 and Wrap 920AR, by measuring the time in seconds needed to walk straight a distance of 20 meters on a smooth pavement without obstacles. We performed this test three times in each patient with and without devices, being the mean of the three PWS measurements considered for the analysis.

We designed three 40-meter circuits of medium difficulty, each one different for each device tested (Wrap and STAR) to avoid a potential learning effect. We measured the time needed by each patient to finish each circuit with obstacles, adding a certain penalization time for each error done during the test, according to the criteria in a penalties table (computed as a compensation for the theoretical gain of time associated to each error done by the patient). This real time for finishing the circuit was used to estimate the walking speed (WS) in each circuit with and without the HMDs. Heart rate (HR) was also recorded before, during and after the test with each device. We compared WS with the PWS of each patient.

We optimized the parameters for the Wrap and STAR models and kept them constant throughout the study: 800 x 600 pixels resolution, equal thickness contours, 80 x 105 image size, black and white, contour by Canny algorithm (0.4), contours color value for each distance (red 78.1 cm, 96.9 cm yellow green, 131.3 cm, blue 220.7 cm), window size 3x3 and scan size 13.

2.3.3. Statistical analysis

We collected all data recorded during the measurements into a database and statistical calculations were performed using commercial statistical software (SPSS version 18.0 for Windows). We calculated the mean and standard deviation for a standard data distribution, while the median and interquartile ranges were calculated in case of non-parametric distributions. For all statistical tests, 2-tailed $p<0.05$ was considered as significant. To compare the results of quantitative variables we used the Wilcoxon test.

3. Results and Discussion

3.1. Depth first-then contour vs contour first-then depth

The comparative analysis of df-tc and cf-tc algorithms in terms of images processed per second rate versus the window size and with sweep values of 7, 10, 13 and 25 (Figure 4) revealed the following:

- In all cases, df-tc algorithm is 3 to 11 times slower than the cf-td depending on the resolution, sweep and window size.
- There is a linear decrease of the processed rate of images with the size of the window applied, higher for higher sweep parameters, as expected. Only the sweep 7 algorithm overcomes the 25 images per second required for an actual real time application video rate.
- Any real-time application should be at least working at 16 frames per second, a rate achieved in all the cases by the cf-td algorithm at 80 x 105 pixels resolution.
- By increasing the resolution to 144 x 192 pixels, the processing times increased even over the double than for 80 x 105 pixels.

Analysing the results of the graphs shown above the conclusion is that the smaller the sweep value and window size, the smaller the computing time. However, decreasing the sweep value or window size does not lead to worse results in the depth calculation (Figure 5).

3.2. Comparative analysis of edge detection methods

A linear decrease of the frames processed per second with the window size was obtained (Figure 6), with higher slope for higher sweep parameters. This decrease seems to depend mainly on the sweep rather than on the algorithm used for edge detection. The comparative analysis among methods revealed that the Canny algorithm was the slowest method, but with a rate over 16 images per second for 80 x 105 pixel resolution in all cases. The differences among algorithms decreased when the window size parameter increased.

A comparison among Prewitt, Sobel, Roberts, and Canny methods is shown in Figure 7. Although the results seem very similar at first sight, they are very sensitive to image noise, with the Canny method being the most accurate of all of them. The images show that the best contour plots are achieved using this method, showing complete edges and low image noise.

Finally, in order to improve the ability of perceiving objects and their positions, and to minimize the influence of color detection skills of the patient, we performed an algorithm to project the information relating to depth/distance by a variation of thickness and color of the edges, as shown in Figure 8.

3.3. Results of the clinical study

The percentage of women included in the sample was 54.55% (n=6), compared to 45.45% of males (n=5). The mean age of subjects was 52.0 ± 23.3 years. The average of the PWS without any device was 1.33 ± 0.28 m/s, 1.14 ± 0.27 m/s with the Star prototype, and 0.96 ± 0.3 m/s with the Wrap device (Figure 9). PWS without device was significantly higher than the value obtained with Wrap (p=0.001) and STAR (p=0.0174) electronic devices. The WS in the circuit without device was 0.76 ± 0.18 m/s, 0.57± 0.18 m/s with the Star prototype and 0.44± 0.12 m/s with the Wrap device. Comparing the measures in the three conditions, the WS without device was significantly higher than that observed with the STAR (p=0.0498) and Wrap electronic prototypes (p=0.002) (Figure 9). Mean differences among PWS and WS were 0.57 ± 0.22 m/s without device, 0.57± 0.20 m/s with the Star prototype and 0.51 ± 0.20 m/s with the Wrap device. The difference among PWS and WS with and without devices was statistically significant (p<0.001).

Without devices, the HR is significantly higher during testing compared to baseline (78.82 ± 15.52 vs. 70.64 ± 10.35 lpm, p=0.006) or the end of the circuit (74.82 ± 11.63 vs. 70.64 ± 10.35 lpm, p=0.004). With the Star prototype, the HR was significantly higher during the test compared to baseline (76.36 ± 16.30 vs. 69.82 ± 16.41 lpm, p=0.01). With Wrap device, the HR was also significantly higher during the test compared to baseline (80.00 ± 14.25 vs. 70.09 ± 9.59 lpm, p=0.01). Comparing the measurements in the 3 conditions, the HR at the beginning of the test was significantly lower with Star device compared to the non-wearing device condition (p=0.01). Mean increase in the HR during the test circuit was 9.45 ± 7.63 lpm without device, 6.55 ± 10.54 lpm with the Star prototype and 4.18 ± 10.15 lpm with the Wrap device.

3.4. Discussion

According to our calculations, the correlation based on the sum of absolute differences between stereoscopic images was the algorithm for the calculation of depth with the lowest computing cost, and for this reason, we selected it for the system developed in the current study. However, any error in the calculation of distances was found to change the shape of the depth map objects regarding real objects when a dense disparity map was used. To solve this problem, we implemented an algorithm to compute the distance from the objects using only their edges, performing the edge detection prior to the disparity computations, a method that also improved the computing times.

Our simulations show that the calculation times were mainly impacted by factors such as the image size, the sweep value and the used window size, but the edge detection method used should be also considered. We have tested different detection algorithms. They were apparently very similar, but Canny algorithm proved the best one in terms of outlining objects and suppressing noise. This was the algorithm that the users preferred to get the best information from the projected image.

Considering an optimized real-time depth estimation algorithm according to our simulations and the Canny edge detection method, a proof of concept system was developed using a commercial HMD and a laptop. With the participation of 11 patients with retinal problems, we performed a preliminary clinical validation. The validation confirmed the usefulness of the system to maintain acceptable levels of PWS and WS during three test circuits, but with lower degree in the change of HR associated to the performance of these walking tasks. This is consistent with the results of other studies reporting improvements in patients' mobility with other HMDs [12,13].

4. Conclusion

We have designed and developed a new concept of HMD with stereoscopic vision using two cameras. The edges of the objects in the images are detected first using a Canny algorithm, the one selected by the users as the best in terms of quality of the processed image. The information of the distances to the objects is processed using a disparity algorithm applied only to the edges of the objects, a method that allows a real-time operation. The processed images show the edges of the objects with information of their distance inserted as color or thickness of the edge. The images are projected in the remaining visual field of the patient. We have shown in a clinical test that the HMD allows the patients' mobility and obstacle detection ability, and reduces their anxiety level. Future research will focus on the improvement of the real-time operation, the portability and the comfort of the system, all of them items demanded by the patients at the end of the trial.

Acknowledgements

We thank the Comunidad Autónoma de Madrid and University Carlos III de Madrid for funding in the Project ATIDiVisTA (Ref. CCG10-UC3M/TIC-4787), and the Comunidad Autónoma de Madrid for the Project FACTOTEM-2 (grant. nº S2009/ESP-1781) and SINFOTON (grant nº S2013/MIT-2790).

There is no conflict of interest of any of the authors of this manuscript.

Authors' contributions: Coco-Martín MB, Pichel-Mouzo M conceptualized and designed the study, coordinated and supervised data collection, carried out the initial analyses, drafted the initial manuscript, reviewed and revised the manuscript and approved the final manuscript as submitted. Torres JC, Vergaz R designed and developed the device, carried out the initial analyses, drafted the initial manuscript, reviewed and revised the manuscript, and approved the final manuscript as submitted. Cuadrado R, Coco RM and Pinto J coordinated and supervised data collection, critically reviewed the manuscript, and approved the final manuscript as submitted.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

References

[1] S. Resnikoff, D. Pascolini, D. Etya'ale, I. Kocur, R. Pararajasegaram, G.P. Pokharel, S.P. Mariotti, Global data on visual impairment in the year 2002, Bull. World Health Organ. 82 (2004) 844–51.

[2] R. W. Massof, Electro-optical head-mounted low-vision enhancement, Pract. Optom. 9 (6), (1998) 214–220.

[3] J.P. Wann, S. Rushton, M. Mon-Williams, Natural problems for stereoscopic depth perception in virtual environments, Vision Res. 35 (1995) 2731–2736.

[4] M. Mon-Williams, J.P. Warm, S. Rushton, Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display, Ophthalmic Physiol. Opt. 13 (1993) 387–391.

[5] E. Peli, The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display, Vision Res. 38 (1998) 2053–2066.

[6] R. Maini, H. Aggarwal, Study and comparison of various image edge detection techniques, Int. J. Image Process. 3 (2010) 1–12.

[7] E. Peli, G. Luo, A. Bowers, N. Rensing, Development and evaluation of vision multiplexing devices for vision impairments, Int. J. Artif. Intell. Tools, 18 (2009) 365–378.

[8] E. Peli, G. Luo, A. Bowers, N. Rensing, Applications of augmented vision head-mounted systems in vision rehabilitation, J. Soc. Inf. Disp. 15 (2007) 1037–1045.

[9] R.C. Bryant, C.M. Lee, R.A. Burstein, E.J. Seibel, Engineering a low-cost wearable low vision aid based on retinal light scanning, SID Symp. Dig. Tech. Pap. 35 (2004) 1540–1543.

[10] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, S. Yaacob, Wearable real-time stereo vision for the visually impaired, Eng. Lett. 14 (2007) 6–14.

[11] T. Limna, P. Tandayya, N. Suvanvorn, Low-cost stereo vision system for supporting the visually impaired's walk, in: 3rd Int. Conv. Rehabil. Eng. Assist. Technol., 2009. Singapore.

[12] M.D. Peláez-Coca, F. Vargas-Martín, S. Mota, J. Díaz, E. Ros-Vidal, A versatile optoelectronic aid for low-vision patients, Ophthalmic Physiol. Opt. 29 (2009) 565–572.

[13] S.L. Hicks, I. Wilson, L. Muhammed, J. Worsfold, S.M. Downes, C. Kennard, A depth-based head-mounted visual display to aid navigation in partially sighted individuals, PLoS One. 8 (2013) e67695.

Figures

Figure 1. Depth estimation: depth first-then contour analysis. Left: Original picture. Right: final result from adding depth information (via color assignation to the objects, with a scale on the right side) to the disparity map. Image resolution 144 x 192 pixels, sweep 27, and window size 9 x 9.

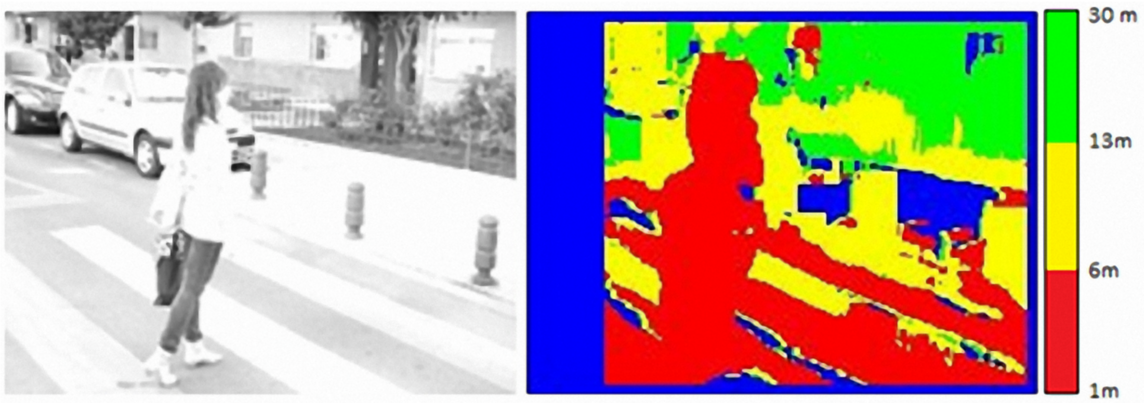


Figure 2. Depth estimation: contour first-then depth analysis. Left: depth map. Right: final result of overlapping depth information to the image. Image resolution 144 x 192 pixels, sweep 27, and window size 9 x 9.

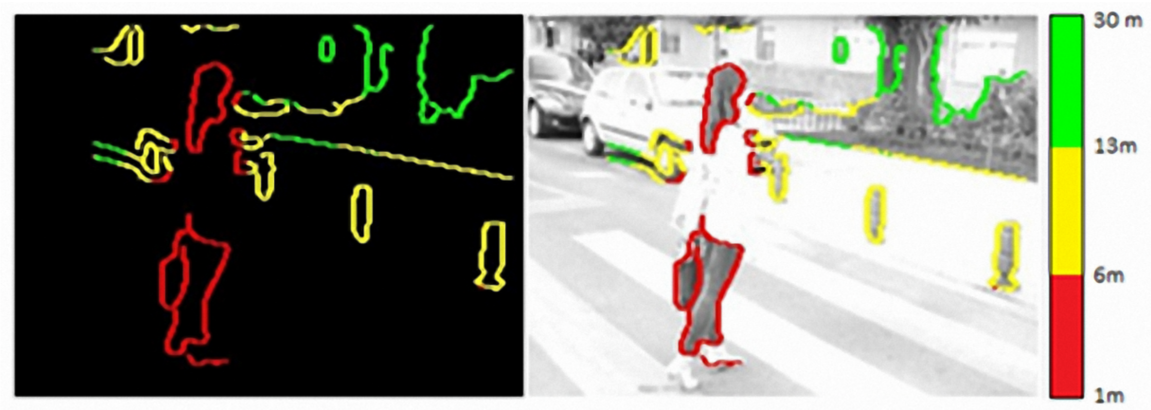
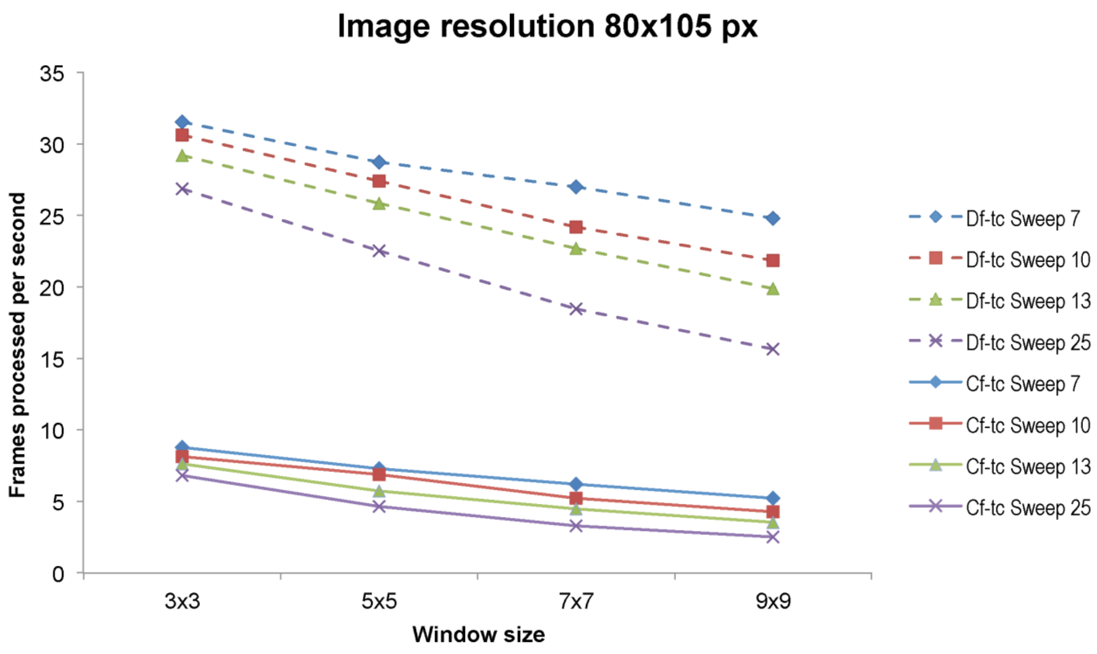


Figure 3. A non-visually disabled person wearing the prototype system. The prototype items are shown in the bottom of figure.

Figure 4. Computing time between traditional and new method for images with different resolution: upper plot, 80 x 105 pixels, and 144 x 192 pixels for the plot below.



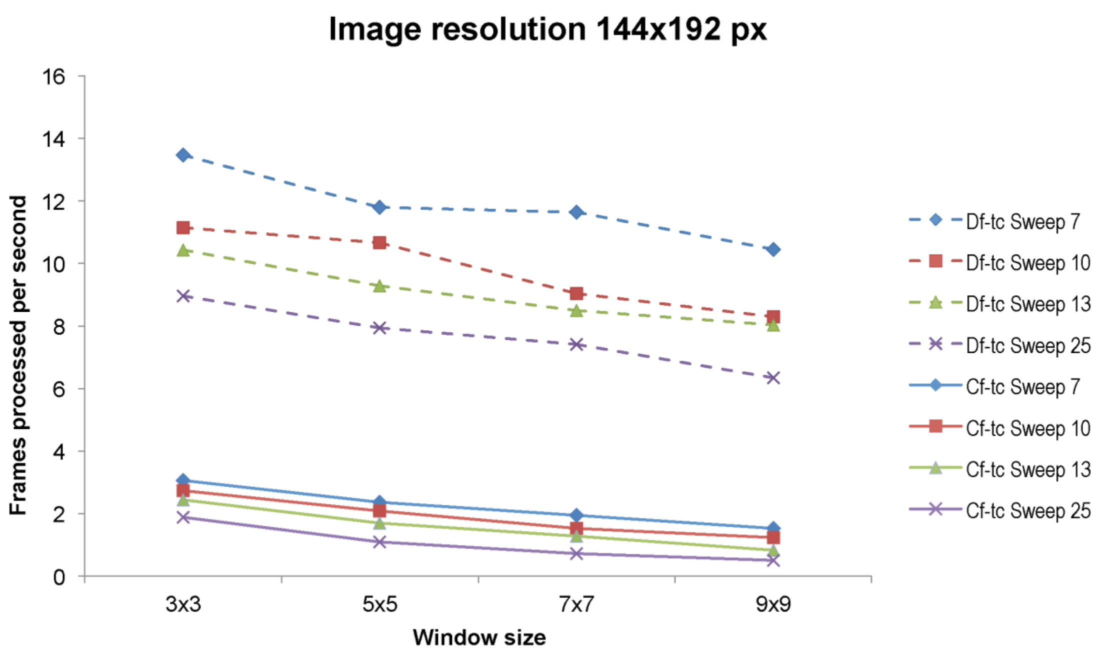


Figure 5. An example of a woman crossing a crosswalk using different calculation parameters: (a) Window size 3 x 3 and Sweep 7, (b) Window size 3 x 3 and Sweep 27, (c) Window size 9 x 9 and Sweep 7, (d) Window size 9 x 9 and Sweep 27 for image resolution of 120 x 160 pixels.

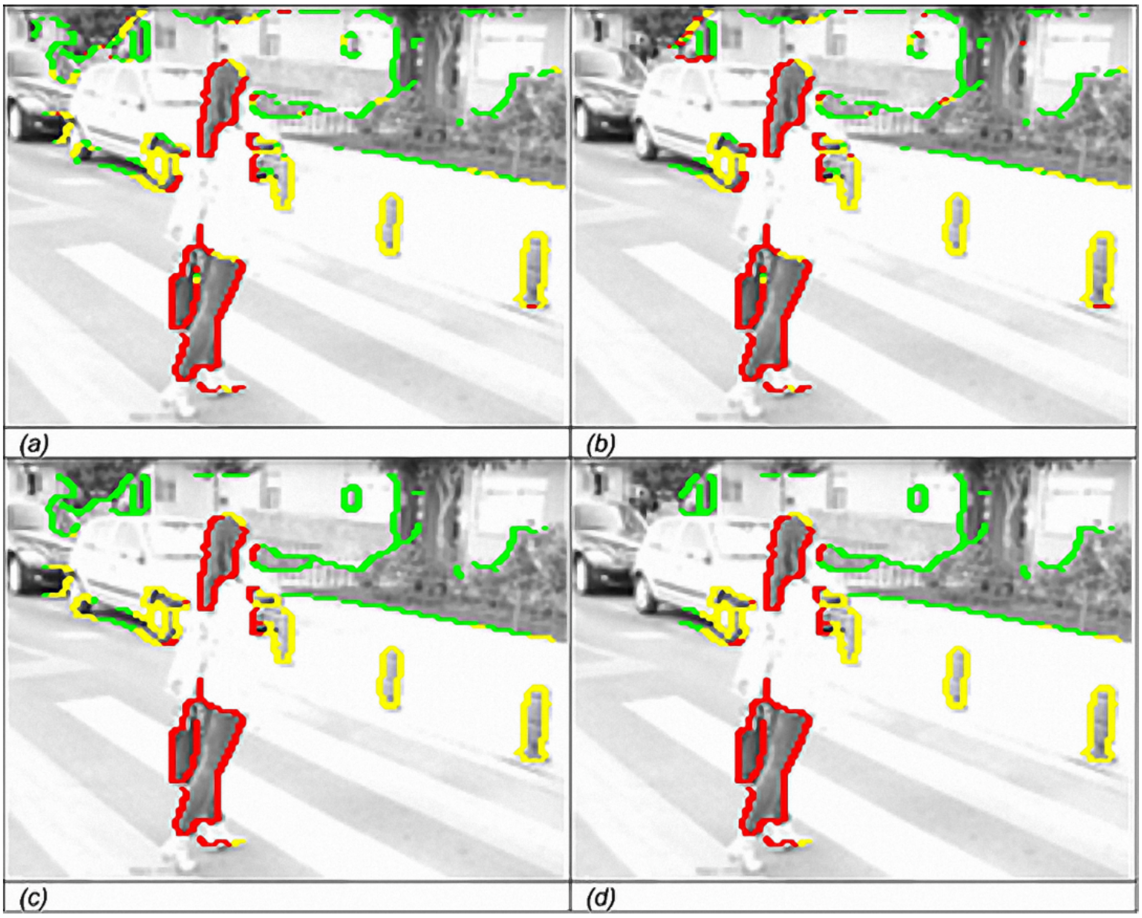


Figure 6. Comparison calculation times by using different detection methods for imaging contours. 80 x 105 pixels in the image, blue means a parameter sweep value of 7, red 13 and green 25.

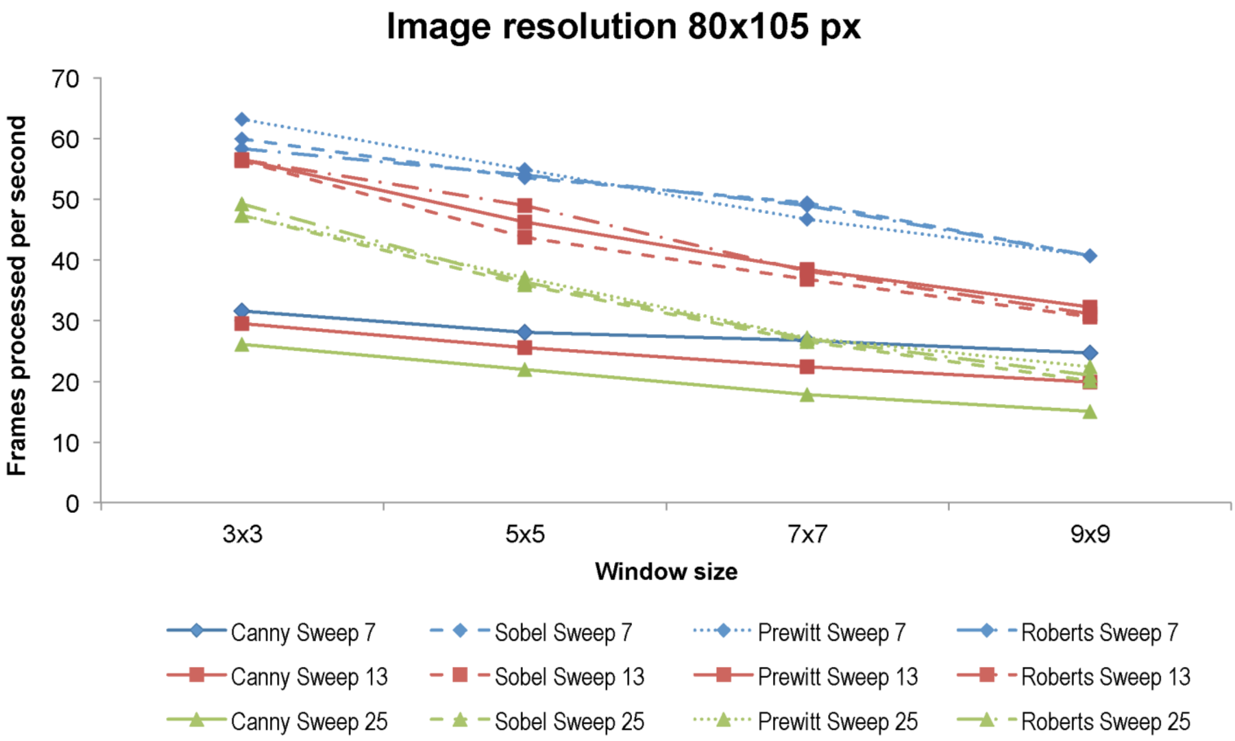


Figure 7. Comparison among different edge detection algorithms: (a) Canny Method, (b) Sobel Method, (c) Prewitt Method and (d) Roberts Method for image resolution of 80 x 105 pixels, sweep 13 and window size 5 x 5.

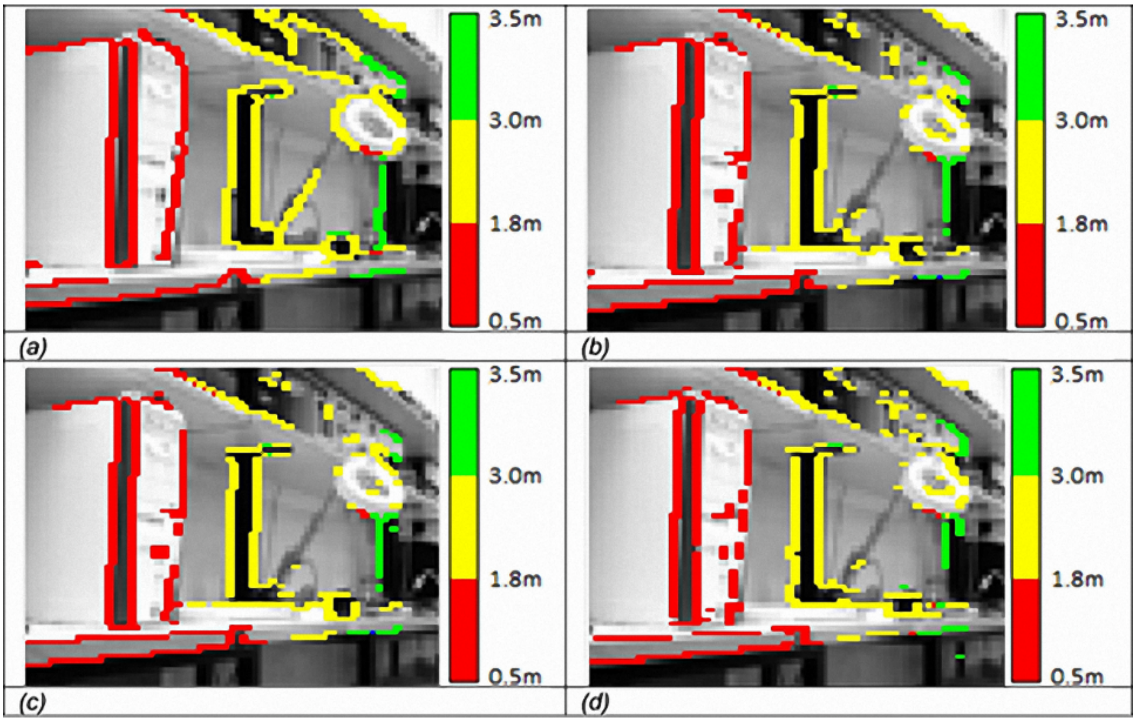


Figure 8. Variation of thickness and color in the edges.

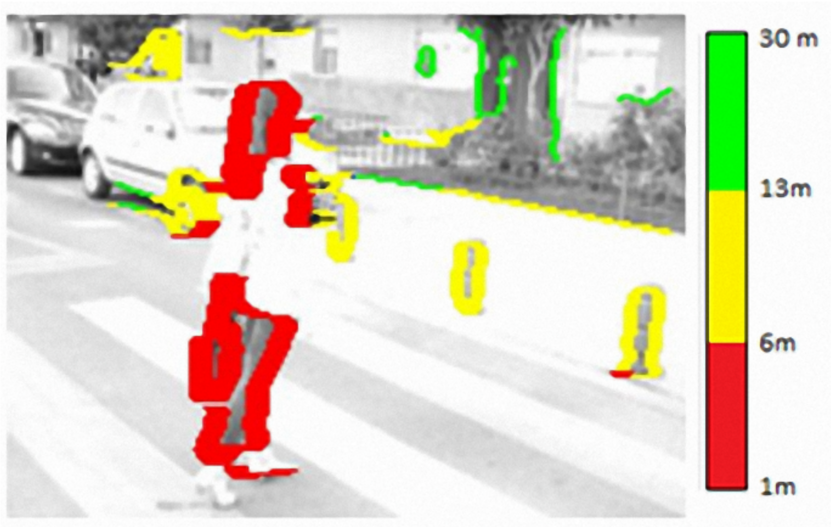
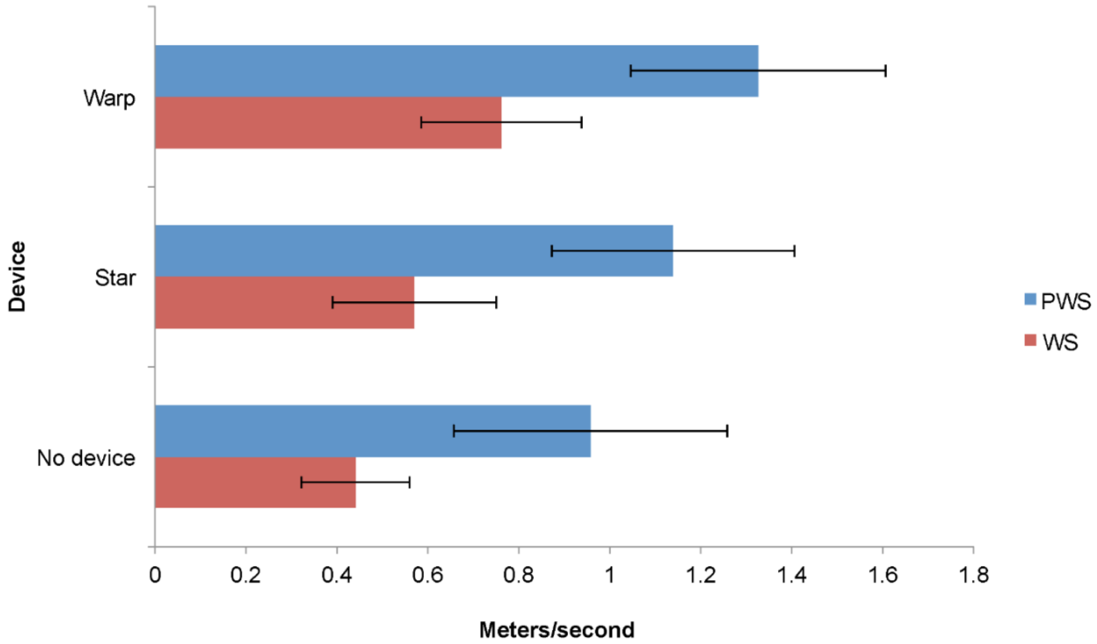


Figure 9. Preferred walking speed (PWS) and walking speed (WS) in the circuits without devices and with the electronic prototypes Star 1200 and Wrap 920AR.



Abbreviations and acronyms: **ARMD** = age-related macular degeneration, **cf-td** = contour first-then depth, **df-tc** = depth first-then contour, **HMD** = head-mounted display, **HR** = heart rate, **PWS** = preferred walking speed, **VA** = visual acuity, **VF** = visual field, **WS** = walking speed.

Development and evaluation of a head-mounted display system based on stereoscopic images and depth algorithms for patients with visual impairment

María Begoña Coco-Martin;[1,3] María Pichel-Mouzo;[1] Juan Carlos Torres;[2] Ricardo Vergaz;[2] Rubén Cuadrado;[1] José Pinto-Fraga;[1] Rosa María Coco;[1]

[1]IOBA Instituto Universitario de Oftalmobiología Aplicada, Universidad de Valladolid, Paseo de Belén, 17, 47011 Valladolid, Spain.

[2]GDAF-UC3M, Grupo de Displays y Aplicaciones Fotónicas, Departamento de Tecnología Electrónica, Universidad Carlos III de Madrid, Av. de la Universidad, 30, 28911 Leganés, Spain.

[3] Research Unit, Hospital Clínico Universitario de Valladolid. Instituto de Ciencias de la Salud de Castilla y León. IECSCYL. Avenida Ramón y Cajal 7, 47011, Valladolid, Spain.

Corresponding author: María B Coco-Martín, PhD. Research Unit. Hospital Clínico Universitario de Valladolid. IECSCYL, Avda. Ramón y Cajal, 47011, Valladolid, Spain. Telephone: +34 983420000 . Email: mbcoco@uemc.es.

Abbreviations and acronyms: **ARMD** = age-related macular degeneration, **cf-td** = first-then depth, **df-tc** = depth first-then, **HMD** = head-mounted display, **HR** = heart rate, **PWS** = preferred walking speed, **VA** = visual acuity, **VF** = visual field, **WS** = walking speed.

Financial Support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions: Coco-Martín MB, Pichel-Mouzo M conceptualized and designed the study, coordinated and supervised data collection, carried out the initial analyses, drafted the initial manuscript, reviewed and revised the manuscript and approved the final manuscript as submitted. Torres JC, Vergaz R designed and developed the device, carried out the initial analyses, drafted the initial manuscript, reviewed and revised the manuscript, and approved the final manuscript as submitted. Cuadrado R, Coco RM and Pinto J coordinated and supervised data collection, critically reviewed the manuscript, and approved the final manuscript as submitted.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.