QUIDEST: REAL-TIME MONOCULAR DEPTH MAP TO AUDIO SIGNAL CONVERSION ALGORITHM

L. Tenze and E. Canessa

The Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy

ABSTRACT

We introduce QuidEst, a simplified computer vision-to-audio signal application aiming to alert any autonomous navigator for potential threats in open spaces. It is based on associations made between real-time depth map computations and spatial audio signals according to the proximity of obstacles. QuidEst is a C-based program that correlates nine specific depth map sub-regions of a video frame to spatial sound effects. The depth map is generated via the MiDaS deep neural network method from a USB webcam or cellular phone camera, and audio threads render the sonification within each sub-region with a combination of faded musical notes. The strength of QuidEst is the minimal and cost-effective hardware requirements for its implementation, together with the software aspects that existing open-source external libraries handle.
QuidEst binaries: https://github.com/canessae/Quidest
Supplemental video: https://www.youtube.com/watch?v=fsVbh53SRio

KEYWORDS

Network Assistive Technologies, Front and Rear Vision, Spatial Sounds, Deep CNN

1. INTRODUCTION

The development of wearable devices to avoid obstacles via stereo vision and audio notifications, best known as electronic travel aids (ETAs), has become a challenging task for almost two decades [1–5]. Lots of these systems integrate different technologies on multiple platforms aiming to benefit specially the visually impaired community. Continued development in navigation-assisted applications are leading to more affordable and effective solutions to improve their quality of life –see, e.g., a comprehensive review in [6].

For example in [3], a simple, portable, hands and ears-free system was proposed to detect obstacles for assisting visually impaired people. This system was based on the use of a 3D camera carried out by the user through indoor and outdoor scenarios. A disparity map was computed live from stereoscopic images to detect potential obstacles in every video frame together with a ground-plane estimation algorithm based on some filtering techniques plus a polar grid representation of the scene. The system also incorporated acoustic feedback to assist individuals while walking forward and approaching obstacles. Different sound frequencies were repeated to inform a blind pedestrian about the presence of obstacles without interrupting the user from hearing other sounds present in the environment. By exploiting smartphones cameras [7,8], mobile devices have also been used to extract a 3D representation of the natural terrain and recognize obstacles by processing images with a ad-hoc Structure from Motion algorithm that takes as input information from the built-in phone gyroscope. These systems consider all structures above the ground-plane as obstacles.

On the other hand, modern deep learning techniques and neural networks with pre-trained models can achieve an accuracy greater than 90% while detecting and recognizing objects in movement [9]. As recently discussed in [10], this opens the possibility to extend these investigations to ETAs by applying deep convolution neural network (CNN) systems in real-time as also adopted in this work. Semantic segmentation using CNN allows to extract meaningful information regarding the features inherent in 2D input images [11].

In light of these ideas, we made an attempt to develop an innovative algorithm that could be used to aid people to ambulate the environment safely and independently, or prevent attacks in dangerous domains. We introduce the C-based program QuidEst[1], which is a fast conversion algorithm that alerts for the presence of obstacles. It is based on associations made between the real-time monocular depth map computations by the MiDaS deep CNN-based method [12], and some spatial audio generated according to the proximity of obstacles. Single video images are captured by any standard (internal or USB) PC webcam or cellular phone camera.

2. PRIOR WORKS AND MOTIVATION

Dynamic object detection techniques can have a wide range of applications including mobile robots, autonomous driving and self-navigation. These techniques mainly focus on detecting obstacles along the ground plane, such as cars, trees, pets, pedestrians, etc. A precise scene understanding to infer an accurately mapping of the surrounding environment is needed in most cases. For some robotic applications, portable fast-moving obstacle experiments using a RGB-D camera have been reported in [13]. The initial studies reported in [1,3] have shown that an obstacle avoidance approach based on the use of a stereoscopic webcam in conjunction with audio feedback can help visually impaired individuals to gain further independence while walking in open, urban scenarios. Post-processing of the RGB and depth images with this approach leads to a better estimation for the obstacle's dimension. The drawback of these systems is that they may be costly or may need extensive computer calculations.

On the other hand, sufficiently simple systems consisting of a single laser beam and a camera to scan the environment have been proposed in [14] using triangulation. They act like a ”virtual” white cane with a larger reach. However, these systems require some time to scan the environment. Horizontal light detection and ranging (LiDAR) has also been used with an accelerometer to investigate the environment in front of a user while walking [15] (see also [10]). The limitation of this approach is that only a slice of the obstacles is estimated at each forward step. An extension of this class of virtual blind cane using a line laser-based vision system, an inertial measurement unit and sound feedback was discussed in [16]. The echolocation is represented by a single tone sound with a magnitude proportional to the distance from the individual to an object. A higher tone indicates that the blind user is closer to the obstacle and only appears when the system detects certain selected obstacles. These systems unfortunately are easily affected by any strong illumination in the surroundings.

Some of these prototypes require processing units not suitable for a compact wearable system. A brief state-of-the-art overview on the use of wearable, lightweight assistive devices for visually impaired people can be found in [17]. Despite the growing assistive technologies for ETAs being developed, more research is needed to overcome the multiple challenges faced by blind users in real life such as reducing constant hand interaction, detecting ground-level obstacles, avoiding extensive training due to complicated alarm codes, and getting precise information about the distances of objects among others. Studies focusing on electronic mobility aids as complementary alternative to a white cane, manifested concrete problems including excessive size and weight,

1 from Latin: 'Quid = What', 'Est = Is' (what is?)

battery life, continuous use and cost complexity, as well as the lack of designs based on user experience and degree of vision deprivation [18].

It is important to mention the motivation behind this work. As a matter of fact, several user studies with the participation of low-vision volunteers (see, e.g., [1,3,5,8]), have proven the effectiveness of ETAs for obstacle detection. This is validated by the fact that blind individuals can still keep using the auditory sense –which is their most important perceptual source, and that they can also keep holding an autonomous white cane without interference. Benefits of the sound of vision devices in outdoor environments as compared to the use of standard white cane have been confirmed in [19]. Participants evaluated positively results consisting in feedback about static and moving objects, surface discontinuities such as staircases, ground holes, joints between walkways and presence of walls. These reliable mobile assistive devices –worn on the head, chest, waist, hands, wrist or feet, collect environmental information and convert it into sound signals or pulses. We implement this type of conversion within a regular webcam Field-of-View (FoV) of 60 − 70o.

3. IMPLEMENTATION

QuidEst is a real-time computer vision-to-audio signal algorithm that associates spatial sounds to specific areas within a video frame in order to indicate the presence of moving obstacles from the surroundings according to its proximity. The minimum hardware needed to test it is any standard PC Computer (to perform the conversion) e.g., Intel Core i5, 64bit and at least 4G RAM, running a recent release of Linux O.S. (23.04LT or newer). The conversion of live images into auditory signals can be done using any camera. Optional hardware that can be used are audio stereo speakers or headphones to listen to the sounds, and an external USB webcam or a smartphone with the free DroidCam app installed to use it as webcam. The latter is freely available from https://www.dev47apps.com. In Fig. 1, we illustrate QuidEst screen output of a phone-as-webcam app connected to a PC Linux.

QuidEst correlates nine specific sub-regions or rectangles of a video frame to fade sound effects to alert for the presence of obstacles. Specifically we apply the MiDaS 2.1 algorithm [12]: i) to process each video frame and produce a reasonable accurate monocular depth map in real-time, and ii) to label somehow the proximity of the obstacles framed around the sub-regions.

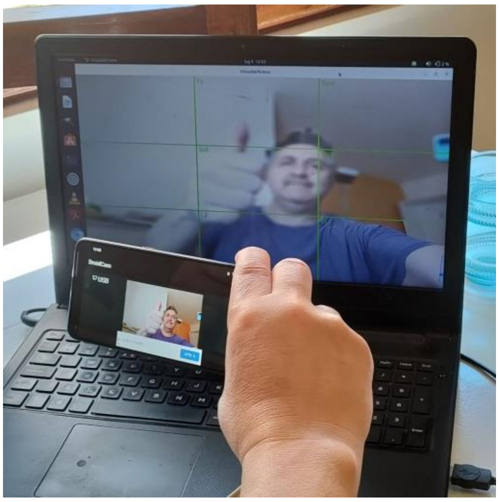

Figure 1. QuidEst in action using Cellphone-as-Webcam.

Our algorithm exploits MiDaS 2.1 CNN which produces a reasonable depth map from a generic single image as input because the available trained network is very effective. Unlike prior works, the depth map generated within QuidEst is fast computed and does not need to be compared against any real disparity map input [20]. This is evaluated by using the Deep Neural Network (DNN) module of the OpenCV library. The DNN module can read the neural network coefficients. It can inference output in an efficient way. The DNN module is configured to prefer CUDA speed-up. The DNN automatically switches to the CPU when the hardware is not present. The MiDaS 2.1 'model-small.onnx' model is adopted to speed up calculations. The analysis of (the nearest white parts of) the 255 gray-scale monocular depth map –divided in depth labels or dm number of layers extracted from a USB webcam or cellular phone camera, are then associated to audio signals. To achieve this the depth map image is first subdivided in nine equilateral rectangles that cover all its pixels or, equivalently, those of the original image as superimposed and displayed in Fig. 2.

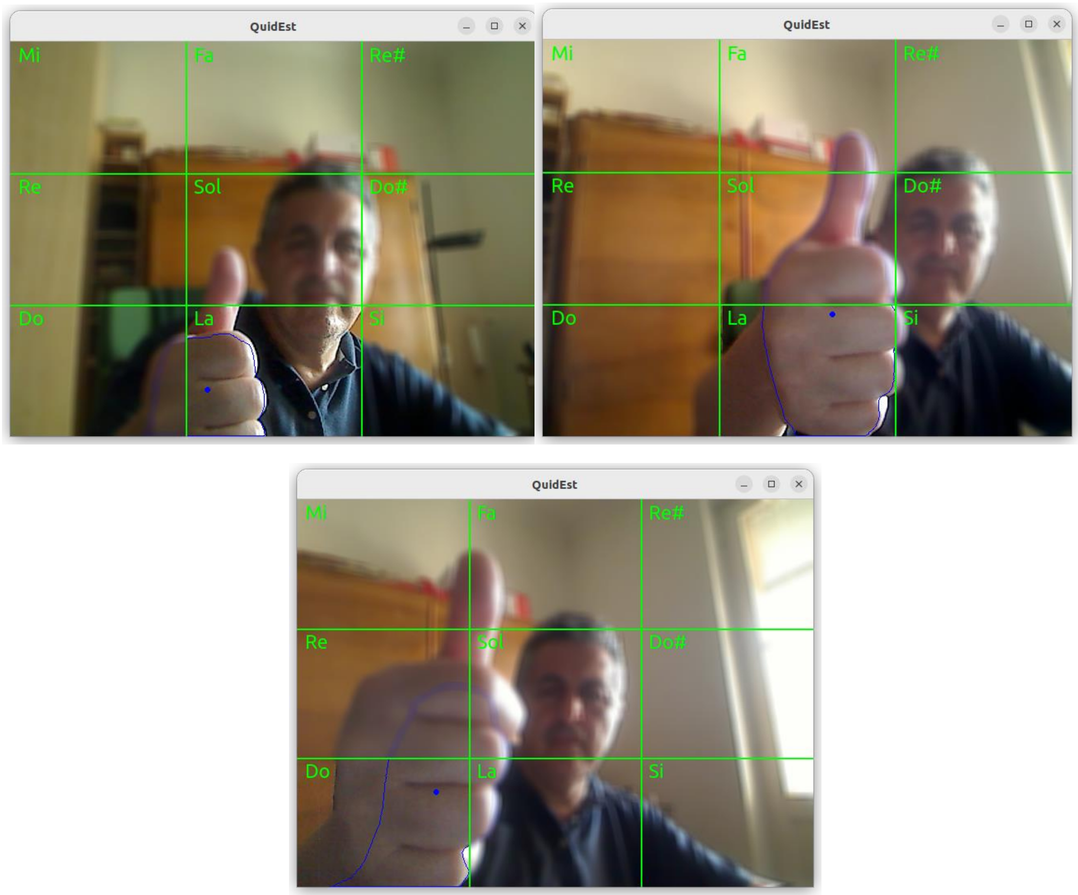

Figure 2. Snapshot of the 9 rectangles covering an input image (or its respective monocular depth map) with associated music notes each as indicated in the corners. Upper Left: ”La” note rings according to the area (blob) being identified in real-time to then fade-out after a few seconds. Upper Right: This ”La” note increases volume since the area being identified has become larger, which in practice, becomes proportional to how nearest (or farthest) the hand appears in the depth map. Center: There is a change of music note to ”Do” and volume as the fist has moved towards the left. This ”Do” note will be heard on the left channel when using stereo speakers or headphones –see video: https://www.youtube.com/watch?v=fsVbh53SRio

In this way, QuidEst associates in real-time nine musical notes to nine specific sub-regions in a video frame in order to send stereo acoustic warnings due to the presence of (front or rear) moving obstacles in the surroundings (depending on the orientation of the camera). The sonification of each rectangle is rendered by audio threads with a combination of simple musical notes. Stereo auditive stimuli are composed by nine faded musical notes such as Do, Re, Mi ⋯ Sharp Re, and ring up only when nearest obstacles are approaching from the surroundings –this is illustrated in the example of figure 2. The 3 different notes on the left of the screen will ring on the left channel of the stereo speakers or headphones. The 3 notes on the center will ring as mono, and the 3 notes on the right of the screen will ring on the right channel. The volume of these notes increases or decreases in relation to one or more areas being identified (blobs), which in practice, are proportional to at how nearest or farthest the objects are approaching the camera. Each of these blobs contain a blue centroid point to assign the corresponding note. The flowchart in Fig. 3 displays all processes involved.

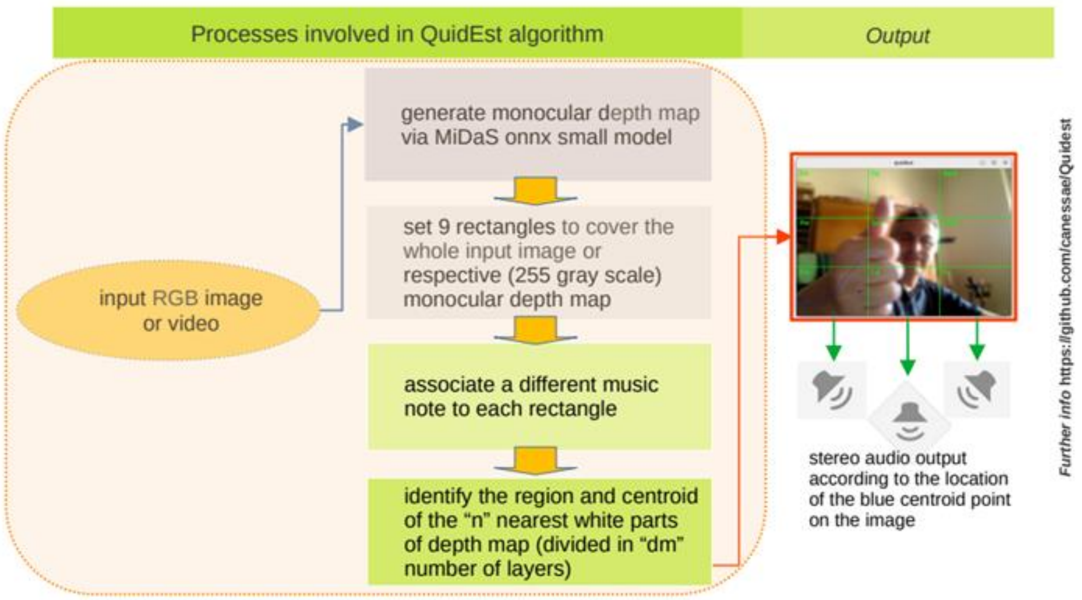

Figure 3. Flowchart of QuidEst for obstacle detection and audio warning.

The amplitude of the tones and the fading in and out effect is implemented in a specific audio class. The in/out fading effect is introduced to avoid cracks in the audio stream. If a new sinusoid comes in or goes out, a crack is produced. To avoid this effect, a proper amplitude control of the sinusoid has been implemented to reduce saturation of the audio wave. The amplitude is changed according to the data provided by the image processing stage. Furthermore, to avoid concurrent access to the data from audio and image processing threads, proper mutexes have been used.

4. DISCUSSION

The idea behind QuidEst is to develop an algorithm that associates spatial sounds to specific blobs within a video frame in order to locate in real-time the presence of (moving) obstacles from the surroundings to then send an alarm according to its proximity. It is mainly based on: i) MiDaS depth map CNN network, ii) Qt library to manage the windows and the audio system iii) the OpenCV library to process images in real-time. the DNN module speeds up the deep neural network inference.

The MiDaS 2.1 is used to produce a depth map from a single image or video frame. QuidEst is able to deal with both MiDaS models the small network and the large one. Good performance is obtained thanks to the DNN OpenCV inference module to grab the image.

For the image processing steps, the image is retrieved from, and elaborated by, a VideoCapture class to feed any other audio and video stages. By exploiting the DNN module, it evaluates the depth map and it applies a threshold to the depth gray-scale image. A threshold function detects the nearest areas of blobs of objects within the processed frame, which is separated from the rest. For each detected blob, its centroid and the area are calculated. The whole image frame is subdivided in 9 regions: top-left, top-center, top-right, center-left, center-center, center-right, bottom-left, bottom-center and bottom-right. The algorithm then loops on all detected blobs to associate each centroid to the proper sub-region. This information is stored and is sent to the audio processing stage to generate unique tones. The amplitude of the audio is changed with respect to the area of each blob. If an object has a large area, the emitted sound will be louder. The last part of the image processing steps manages a blurring effect inside the sub-regions: if a centroid (or more than one) is detected inside a sub-region, the image in that sub-region is normal. Otherwise, if no centroid is present, the sub-region is blurred. This way, the user can visually detect where centroids are present or not if needed. The areas on the screen around a blue centroid correspond to the nearest objects. Instead, the farthest objects will appear blurred and will not deliver any sound. Within QuidEst it is possible to invert the video image on the screen, and related sounds, and make these specular. This allows to adapt QuidEst as a tool to generate alerts related to the detection of rear-view objects while walking forward.

The audio device is controlled by exploiting the Qt QAudioOutput, QAudioDeviceInfo and QIODevice classes. The Qt classes are properly included in an audio context class and it is used in a thread separated by the main GUI thread. The thread is refreshed every 20 ms in order to check if something new has happened and if some new objects have been detected from the graphical processing stage. The audio thread feeds the audio buffer with the proper combination of musical notes/tones which will be rendered by the real audio output. The buffer is filled in by sinusoid of notes as described next.

The MyNotes class defined in QuidEst declares the frequencies associated with every note. When a blob enters a given region, the application emits a sound associated with that particular region. As mentioned, the right regions produce sounds only in the right channel, the left ones in the left channel and the middle regions produce a mono signal (left = right). When a note needs to be added to the output signal, the system creates a new sinusoid pure tone with a proper amplitude, which depends on some parameters:

1. the amplitude is associated with the size of the blob inside the region of detection,
2. the amplitude needs to be faded out when the blob is static for more than 3 seconds (default value).
3. In order to avoid cracks when the note is added and when it is removed from the output signal, a fade-in fade-out procedure is implemented.

The most critical part is the management of point 3. To implement the fade effect a linear fading has been adopted. For example, the output signal at the initial time t may be composed by a single frequency:

$$S_{output}(t) = A_i(p1, p2) \times sin(2\pi f_i \times t), \qquad (1)$$

where the function Ai is the current amplitude of i-th generated sinusoid. This is a function of p1 and p2 relating the items 1 and 2 mentioned above. The amplitude Ai is a function of: i) i –the i-th sinusoid (in the current version of QuidEst there is a maximum number of 9 sinusoids), ii) p1 – the parameter which increases or decreases the amplitude of the sinusoid, iii) p2 –the parameter that zeroes the amplitude when the sound is emitted for more than 3 seconds.

Fading is then implemented with another factor. If a new blob appears, a new sinusoid is added as follows

$$S_{output}(t) = A_i(p1, p2) \times sin(2\pi f_i \times t) + F_j(t_{start}) \times A_j(p1, p2) \times sin(2\pi f_j \times t), \quad (2)$$

where Fj (tstart) is the function implementing the fade-in effect for the note j. The amplitude of the note is modulated by this new function in order to progressively introduce (from tstart) the new note while avoiding annoying cracks in the output audio sound.

A clear novelty and originality of our QuidEst software is its light-weight nature due to its simple output format. The strength of the present method is the minimal and cost effective hardware requirements for its implementation together with the software aspects which are handled by existing open source external libraries and algorithms (namely, OpenCV, Qt, MiDaS). The minimal requirement to test QuidEst software is a standard PC with 4 GB RAM whose O.S. is Linux. The kernels used are for 22.04LT and 23.04LT for fast run. QuidEst accepts an image/frame as input, which is processed in real time by a CNN to detect obstacles in the observed space, and finally the information is returned to the user as a sound notes to alert potential dangers nearby (in front or rear).

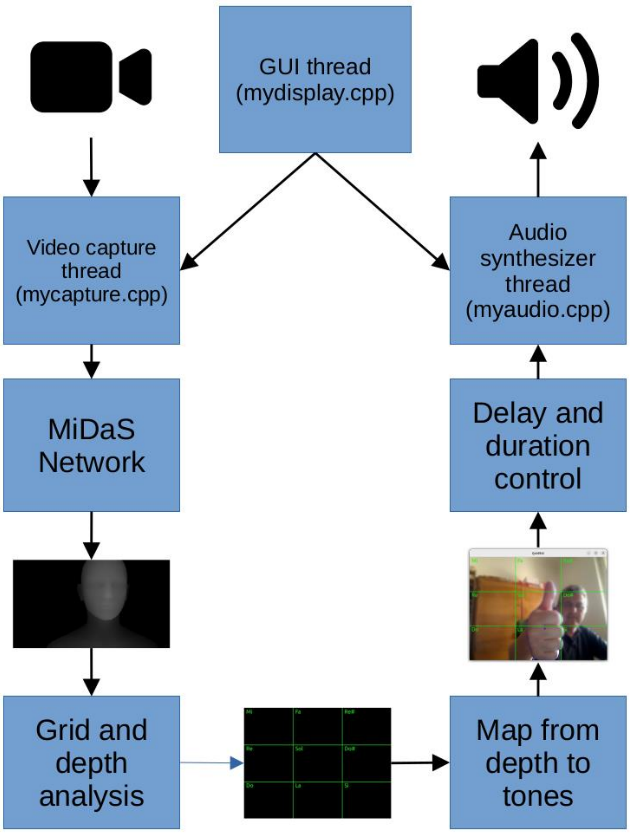

Figure 4. QuidEst simple framework.

The RAM needs depends, for example, on the accuracy of the real-time display outcome by the spatial resolution of the images or frames and the MiDaS analysis of the 255 gray scale depth map being divided in given number of layers (20 by default), which is extracted from a single USB webcam or cellular phone. The performance of the MiDaS algorithm has been largely discussed in [12]. The software considers nine (or more configurable) subdomains and map a different spatial sound to each region to inform about the direction of the detected nearest object. In fact, the most inspiring feature of QuidEst is that it can be used as a front and/or rear warning single webcam-based guidance system while other packages cannot manage both possibilities.

At present only the nearest obstacle is alerted by sounds as indicated in blue contours in Figure 2. The case of multiple neighbour obstacles approaching at the same time can be improved by making all sounds to ring (almost) simultaneously time whose volumes are properly weighted based on the distance in each screen subdomain, which are configurable (nine by default). It could also be extended via an App and other applications to increase the number of depth map sub-regions to become essentially a virtual Theremin [21]. The sounds (Do, Re, Mi $\cdots$) used to indicate far/near positions are also configurable (by default we used the simplest music scale). QuidEst could also incorporate an extra zooming feature to better highlight the sub-regions when approaching objects.

This novel class of depth map to sonar-like conversion system with a sonification within each sub-region rendered by audio threads with a combination of faded musical notes given by Eq.(2) and illustrated in the Flowcharts in Figs. 3 and 4, may be useful as a single webcam-based guidance system. Monocular depth map MiDaS is known to achieve fast detection accuracy at typical detection speeds of 32fps or more. To some extent an implementation of QuidEst may be also useful for rear vision awareness for all. The performance of QuidEst in terms of detection and processing speed shows encouraging results as it stands –see YouTube example video: https://www.youtube.com/watch?v=fsVbh53SRio

5. CONCLUSIONS

In this work we introduced QuidEst –our new application aiming to alert any autonomous navigator for potential threats along open (rear or frontal) spaces. We aim to alert any autonomous navigator for potential threats along open spaces, rather than to improve greatly on mapping the environment. By construction, QuidEst may help people to find safer frontal paths or prevent rear attacks through unknown environments. Acoustic signals warn about the relative obstacle position and its proximity. The main difference of QuidEst with respect to many other vision-to-audio signal approaches is that this is a simple alternative for real-time object detection for ETAs based on the fast conversion of monocular depth map to specific acoustic signals. As stated in Section 1, the idea to process visual information and output it as audio signals has been attempted extensively in the community. This is an interesting subject which requires further investigations because there is still not a definitive low-cost solution available.

QuidEst can also be embedded in a Raspberry Pi with Alpine Linux installed. In addition, QuidEst can also be adapted to mobile LiDAR devices to improve on the processing algorithms, avoiding the use of MiDaS convertion, to offer easier wearability. Currently mobile devices are powerful even if the proposed CNN network could be difficult to run on a mobile hardware. The most promising solution could be to exploit the internal depth camera, becoming popular in latest modern devices. In this case the depth map could be obtained by a direct access to the depth camera while avoiding extra computational costs. Moreover the choice of libraries such as openCV would simplify the porting and the implementation into Android devices: openCV can be easily integrated as a low-level library in an Android application: the C source code of QuidEst provides good performance with respect to a Java implementation. In all cases, the use of a single webcam limits the $(60 - 70o)$ FoV to label nearest obstacles moving from around 2 meters towards a user. In addition, QuidEst could easily replace the sound signals with, e.g., vibrations or the switching of RGB leds. All these directions have the potential to be explored.

QUIDEST BINARIES

Further information, binaries, papers, presentations, manuals, or to report bugs, can be found at https://github.com/canessae/Quidest

REFERENCES

[1] Saez Martinez J.M., Escolano Ruiz F., ”Stereo-based Aerial Obstacle Detection for the Visually Impaired”, In Proc. Workshop Computer Vision Applications for the Visually Impaired, Marseille, France, 18 Oct 2008. *HAL open access archive:* https://inria.hal.science/inria 00325455 (Last visited 24/6/2024)

[2] Brown D., Macpherson, Ward J., ”Seeing with Sound? Exploring Different Characteristics of a Visual-to-Auditory Sensory Substitution Device”, *Perception* 40 (2011) 1120. doi: 10.1068/p6952

[3] Rodriguez A., Javier Yebes J. et al., ”Assisting the Visually Impaired: Obstacle Detection and Warning System by Acoustic Feedback”, *Sensors* 12 (2012) 17476. doi: 10.3390/s121217476

[4] Hu W., Wang K. et al., ”A Comparative Study in Real-Time Scene Sonification for Visually Impaired People”, *Sensors* 20 (2020) 3222. doi: 10.3390/s20113222

[5] Messaoudi M.D., Menelas B-A.J., Mcheick H., ”Review of Navigation Assistive Tools and Technologies for the Visually Impaired”, *Sensors* 22 (2022) 7888. doi: 10.3390/s22207888

[6] Isazade V., ”Advancement in Navigation Technologies and Their Potential for the Visually Impaired: A Comprehensive Review”, *Spatial Information Res.* (2023) https://link.springer.com/article/10.1007/s41324-023-00522-4 (Last visited 24/6/2024) doi: 10.1007/s41324-023-00522-4

[7] Caldini A., Fanfani M., Colombo C., ”Smartphone-based Obstacle Detection for the Visually Impaired”, in: Murino V., Puppo E. (eds) Image Analysis and Process. ICIAP 2015. *Lecture Notes in Computer Science*, 9279. Springer, Cham. Doi: 10.1007/978-3-319-23231-7 43

[8] Neugebauer A., Rifai K. et al., ”Navigation Aid for Blind Persons by Visual-to-Auditory Sensory Substitution: A Pilot Study”, *PLOS ONE* 15 (2020) e0237344. doi: 10.1371/journal.pone.0237344

[9] ”Design and Implementation of Real-Time Object Detection System based on Single-Shot Detector and OpenCV”, Wahab F., Ullah I., Shah A. et al., *Front. Psychol.* 13 (2022) 1039645. doi: 10.3389/fpsyg.2022.1039645

[10] Cheng X., Qiu S. et al., ”Understanding Depth Map Progressively: Adaptive Distance Interval Separation for Monocular 3d Object Detection”, *ArXiv*: https://arxiv.org/abs/2306.10921 (Last visited 24/6/2024) doi: 10.48550/arXiv.2306.10921

[11] Malukas U., Maskeliunas R. et al., ”Real Time Path Finding for Assisted Living Using Deep Learning”, *J. Universal Comp. Sci.* 24 (2018) 475. doi: 10.3217/jucs-024-04-0475

[12] Ranftl R., Lasinger K. et al., ”Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-shot Cross-dataset Transfer”, IEEE Trans. Pattern Analysis. *Mach. Intell.* 44 (2022) 1623-1637. doi: 10.1109/TPAMI.2020.3019967 ArXiv: https://arxiv.org/abs/1907.01341v3 (Last visited 24/6/2024) doi: 10.48550/arXiv.1907.01341

[13] Saha A., Dhara B.Ch. et al., ”Efficient Obstacle Detection and Tracking Using RGB-D Sensor Data in Dynamic Environments for Robotic Applications”, *Sensors* 22 (2022) 22. doi: 10.3390/s22176537

[14] Yuan D., Manduchi R., ”Dynamic Environment Exploration Using a Virtual White Cane”, In Proc. IEEE Comp. Soc. Conf. on *Computer Vision and Pattern Recognition*, San Diego, CA, USA, June 2005 (1) pp. 243–249. doi: 10.1109/CVPR.2005.136

[15] Palleja T., Tresanchez M. et al., ”Bioinspired Electronic White Cane Implementation Based on a Lidar, A Tri-axial Accelerometer and a Tactile belt”, *Sensors* 10 (2010) 11322. doi: 10.3390/s101211322

[16] Dang Q.K., Chee Y. et al., ”A Virtual Blind Cane Using a Line Laser-Based Vision System and an Inertial Measurement Unit”, *Sensors* 16 (2016) 95. doi: 10.3390/s16010095

[17] Prada A., Forero S., ”A Belt-like Assistive Device for Visually Impaired People: Toward a More Collaborative Approach”, *Cogent Engineering* 9 (2022) 2048440. doi: 10.1080/23311916.2022.2048440

[18] Abreu D., Suarez A. et al., ”Safe Displacements Device for All Conditions Blind People”,*Electronics* 12 (2023) 2171. doi: 10.3390/electronics121021718

[19] Zvoristeanu O., Caraiman S. et al., "Sensory Substitution for the Visually Impaired: A Study on the Usability of the Sound of Vision System in Outdoor Environments", *Electronics* 10 (2021) 1619. doi: 10.3390/electronics10141619

[20] Leong K.T., Lim S.M., "SurDis: A Surface Discontinuity Dataset for Wearable Technology to Assist Blind Navigation in Urban Environments", 36th Conf. Neural Inf. *Processing Sys. Datasets and Benchmarks Track* (NeurIPS 2022), New Orleans. ISBN: 9781713871088 https://proceedings.neurips.cc/paper_files/paper/2022/hash/ 988120df77d6c767995febd7ff616517-Abstract-Datasets_and_Benchmarks.html (Last visited 24/6/2024)

[21] Usabaev B., Eschenbacher A. et al., "The Virtual Theremin: Designing an Interactive Digital Music Instrument for Film Scene Scoring", *i-com* 21 (2022) 109. doi: 10.1515/icom-2022-0007

AUTHORS

Livio Tenze (Ph.D in Electrical Engineering) and Enrique Canessa (Ph.D in Physics) are with the Science, Technology and Innovation (STI) Unit of the ICTP in Trieste, Italy.