Full Paper

NNGT Int. J. on Information Security, Vol. 3, Jun 2015

camera - based indoor neural network for visually impaired people "Patented"

Professor: M. Tariq Iqbal

Department of Electrical and Computer Engineering
Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St. John's, Newfoundland, Canada, A1B 3X5
Office: EN-3062: E-mail: tariq@mun.ca : Phone: (709) 864-8934

Dr. Abdalla Ghait Artaime

Department of Electrical and Computer Engineering
Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St. John's, Newfoundland, Canada
Office: EN-3052: E-mail: aa8231@mun.ca : Phone: (709) 753-6389

Abstract — in this paper I will develop a device to help visually impaired people find way around in indoor/outdoor unknown environment. The system is based mainly on obstacles detection in a video stream. Using video/image processing for obstacles detection and pathways determination needs less sensors and smaller than other approaches like using sticks and sensors. Using video allow us utilize smart computer vision techniques to detects and information than other approaches. The system will consist of three parts: camera, video processor, and speech generator. The video processor will process the video stream, detect objects, obstacles and pathways, then deliver the extracted information to the speech generator to generate suitable phrases to inform the user about his/her path. I propose a low cost system that can detect obstacles in user's path while walking in indoor or outdoor environment then send speech alert to user to help navigate him. In general our system will consist of camera to get frames and a microprocessor to process those frames, finally send the decision as sound alert to the user whether there are obstacles in the path or not.

Keywords — blind people, openCV, GPS, RoboRealm

I. Introduction

Historically, visually-impaired people used a white cane (international symbol of blindness), to navigate safely and avoid obstacles in their path. Others employed trained dogs to help them navigate safely outdoors and indoors. Some may not use any type of tool for guidance, and just depend on their memory of well-known places and their senses to detect and avoid a probable obstacle. Recently, technical advancement provided a third option for visually-impaired people through the development of electronic systems to detect obstacles and help visually-impaired people avoid them. Visually-impaired people need a system to provide them with information about surrounding environment, so they use this information to make decisions on their navigation path and their movement. These systems could be simple like sticks and could be complex like video camera, in our proposal I suggest a low cost, comfortable wearable and efficient system that uses a video camera to help visually-impaired people navigate safely in indoor unknown environment. User will wear a camera to explore the surrounding area the system will alerts them to navigate safely.

Such systems that detect obstacles are currently deployed in automotive systems and robots to help cars/robots navigate safely and independently. Automatic navigation systems for cars and robots are more complex and require a higher level of intelligence than systems used by visually-impaired people. The following section will discuss a group of systems developed for visually-impaired people safe navigation as well as systems developed for robotic and automotive systems navigation.

II. LITERATURE REVIEW

Object detection is an important field for many applications of computer vision, image and video processing. However the best accuracy and fast invariant detection function under changing object states such as (position, scale, illumination and noise) is a central aspect problem of the object detection in the video frames and images that cannot be realized by using sequential processing with General Purpose Central Processing Unit processor. To speed up the highly intensive calculation required, the implementation operation is achieved by using parallel processing.

Active sensors based approaches: Huge work has been done in this field, these systems use one or more sensors, the most famous sensors are infrared and ultrasonic. The basic idea behind these sensors is that they send a wave "signal" forward and receive it back due to collisions with obstacles, the time between sending and receiving the signal is measured then used to determine if there are obstacles in the path and the distance needed to reach them. Such systems are simple and could give good results in a normal environments figure 2.1 shows an example of these systems. Figure 1: Body attached sensor [1]

Figure 1: Body attached sensor [1]

DOI : 03.IJIS.2015.1.4

GPS based approaches: GPS systems use signals to find the absolute location, based on this location they navigate the user toward the aim, there is a big problem when GPS is used in the area where obstacles such as buildings, trees and bridges obstruct sky and interfere with reception of GPS signals. Then the GPS becomes less accurate, and in enclosed areas, it becomes unusable, because it has no signal. To counteract this, there are several technologies that are used to determine the position within an enclosed area like GSM, WLAN, GPS based system could be integrated with another type of systems like Ultrasonic to add navigation features into them, for example Ljupko Šimunović [2] Suggested a system that can determine the location of user using GPS and use Ultrasonic to detect obstacles in the path, then based on the location and detected obstacles it navigate the user toward his aim.

Vision based approaches: Camera based approaches depend on the number of cameras used and 2D or 3D structure for images and frames.

. Multi-cameras versus single camera: The most common approach for obstacle detection and avoidance is the use of multiple cameras. Thus for example, David Coombs and Karen Roberts [3] propose two cameras looking obliquely to steer between objects. The left and right proximities have been compared to steer through the gap. Another similar development is a vision system capable of guiding a robot through corridor-like environments by Argyrols and Bergheim [4]. It uses three cameras, one for central forward vision and the other two for peripheral vision. Analogous approaches have been proposed and successfully applied for various robotic platforms. Representative examples are Ref [5] for Stereo Vision (most common for ground robots) and Ref [6] for fusing Radar and Vision for obstacle avoidance on cars.[9] Single camera approaches also used to detect obstacles for example Randal C Nelson[7] proposes the use of certain measures of flow field divergence as a qualitative cue for obstacle avoidance. It has been shown that directional divergence of the 2D motion field indicates the presence of obstacles in the visual field of an observer, undergoing generalized rotational and translational motion. Also In their paper [8], Young et. al. presents an approach to obstacle detection, using optical flow. A linear relationship, plotted as a line called reference flow line, has been used to detect discrete obstacles above or below the reference terrain.

.2D versus 3D: Solving a 2D problem of obstacle detection and avoidance is simpler than solving a 3D problem. The 2D problem deals with the intensity map at each pixel on an image, in which obstacles need to be avoided. The 2D obstacle detection hence generally solves only the problem of “directions to avoid” and need not generate a scene model. For the specific case of obstacle avoidance, when we only need to know the regions of its way ahead that are occupied by obstacles. No information like shape of the objects, their absolute positioning in the world or the understanding of the relationships between these objects is required. Consequently the image data may be directly used without a reconstruction of the 3D world of motion. Therefore, no explicit knowledge about the camera parameters and camera-to-ground coordinate transformations is required. On the contrary, a general 3D obstacle detection problem solves for all such attributes of world. Vision-based approaches are an active field of researches, huge work were done here and the wheel is still rolling, in general systems with different algorithms were built and succeeded to give the desired output so we can't easily say this system is better than this system or this system is more accurate because each one was built for a certain conditions and to solve specific kind of problem.

Figure 2: Example of vision based systems

III. METHODOLOGY

Because I have curiosity to take my proposal into another level I propose a vision based approach system. I propose a wearable low cost system that can detect obstacles in user’s path while walking in indoor environment then send speech alert to user to navigate him. In general our system will consist of camera to get frames and a microprocessor (PC) to process these frames, finally send the decision as sound alert to the user whether there are obstacles in the path or not, Figure 3 describes the general system structure**.**

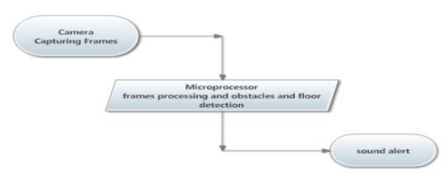

Figure 3: General system structure

Obstacle Detection

Mixing the terms objects and obstacles detection is a common mistake because they are totally different. In general, obstacles detection is the process in which the system output indicates if there is an open path or obstacles on the way. In my proposal I deal with two kinds of obstacles, stationary and moving obstacles we detect them then output this detection for the user. On the other hand, object detection is the process in which the output indicates where a predefined object is detected or not, in other words objects detection is the process of searching for specific features in the images or frames to detect if a certain object/objects is/are there or not. In my system I need

DOI : 04.IJIS.2015.1.4

obstacles detection not object detection since I am not interested to know exactly what the object is; my interest is the existence of it. Obstacles detection is used in many fields like cars and robots, it has many useful applications in their moving and navigation, recently obstacles detection is used in applications that help visually impaired people in their moving and navigation.

Obstacle Detection Approaches

Obstacle detection and avoidance has been an active area of research in the recent past. Some researches were done using some kind of active sensors, e.g. laser range finders, sonar, radar, infrared and ultrasonic. Also passive camera based research has been done, either using stereo vision (multiple cameras) or a single camera. [9] Figure 4; illustrate approaches used in obstacles detection based on the technology they use.

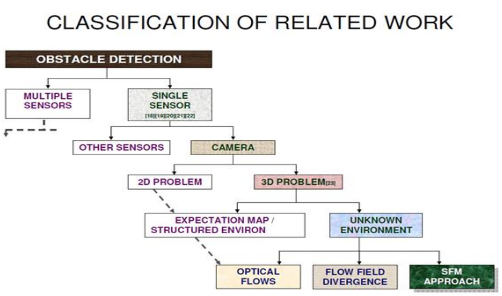

Figure 4 :Classification of related work [9]

Many obstacle detection methods have been proposed, but none offer a clear advantage over the others. As mentioned before, vision-based systems and obstacle detection methods can be divided into two categories; those using stereo vision and those using a single camera. Well calibrated stereo systems can achieve impressive results, but at a severe cost. Namely, the baselines for these systems are usually greater than one meter to enhance depth resolution. They require high quality cameras with long focal-length lenses for adequate spatial resolution. Also the best results have been achieved with three cameras, rather than just two. These requirements restrict the adoption of stereo systems to market products. We conclude that Stereovision systems have the advantage of instantaneous depth perception so it makes use of the disparity field to infer the depth of image features. This allows detection of the floor (if it is unknown) and detection of obstacles as any feature off the floor. However Stereo-vision methods require two cameras and a proper calibration between them in order to function. It needs more expensive (both computationally and financially) and require more maintenance than monocular systems. Monocular systems operate at a reduced cost and require little maintenance, but lack the depth perception of stereo systems. Monocular systems can be divided into three categories:

- Appearance-based methods.
- Motion-based methods.
- Reconstruction-based methods

Appearance-based methods work under the assumption that obstacles appear differently from the surface. So detection is based on color, texture or shape. These methods have low computational cost but are easily fooled in everyday scenarios. [10] for example Ulrich and Nourbakhsh. [11] Use only color information to detect obstacles for the purpose of robot navigation. They classify each pixel as ground or obstacle based on how similar the pixel's color is to the ground. To this end, a hue and intensity histogram is built from a region near the bottom of the image assumed to be ground. If a given pixel color is below a threshold in either of these histograms, it is considered an obstacle. The ground color histograms are updated over time, based on information coming from the camera. The strength of appearance-based methods is speed, simplicity and the ability to detect very small obstacles. They suffer from underlying assumptions. First, obstacle detection can only occur if obstacles differ in appearance from the ground. Second, obstacles cannot occupy the area near the vehicle assumed to be ground; otherwise their appearance gets incorporated into the ground model. Motion-based methods compare two or more images of the video stream and detect obstacles by virtue of their motion is differing from the grounds. So it ignores color and shape but rely heavily on the motion of image features and optical flow. These methods are more robust than appearance-based methods since they detect on the basis of 3D structure rather than appearance. [10] Motion-based methods generally work in the following way. The ground, assumed to be planar, is expected to move a certain way on the image. A parametric model is constructed to capture the essence of this motion. Then, the floor motion is observed, and the parameters of the motion model are estimated. Regions of the image that agree with this model are considered floor, and the rest are considered as obstacles. Reconstructionbased methods or Structure-based methods, work by explicitly reconstructing the 3D scene from motion. Obstacles are detected from image features that lie above the floor in the 3D model. Reconstructionbased methods are more complex and computationally expensive than the previous two categories, but offer a direct estimate of the obstacle's location in space. Structure-based methods detect obstacles by an explicit 3D reconstruction of the scene. For this to be possible with a single camera, the camera must be in motion. The camera motion creates a sequence of images of the scene taken from different positions. If the scene is static, then there is no distinction between a set of images taken from a single camera at different times. The main tasks in most structurebased methods are motion estimation, 3D scene reconstruction, establishing the ground plane and detecting obstacles as features above the ground plane. [10]

Algorithm

Both color and texture are used as appearance-based methods that work under the assumption that obstacles appear differently from the surface, using color only as a classifier will fail when the sample images were taken under bad illumination condition and when there were obstacles with very similar color appearance, so using it

DOI : 03.IJIS.2015.1.4

alone will not be efficient, also if we use texture alone as a classifier good result will be obtained with high computation cost, so I propose an algorithm in which we use both of them as a classifier so I mix the simplicity of color features and the accuracy of texture features, next I will discuss this algorithm in details. Algorithm will generally consist of two parts: a supervised learning step and a classifier step. In the learning step the algorithm learns the floor feature. The classifier then compares a patch of pixels in new images to the learned model and decides if the patch is a path or an obstacle. I have work in two directions to implement my system, one is using OpenCV and the other is using RoboRealm.

OpenCV algorithm

OpenCV (Open Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision, developed by Intel. The library is cross-platform. It focuses mainly on real-time image processing. As mentioned above OpenCV is cross platform so we can use any programming language.

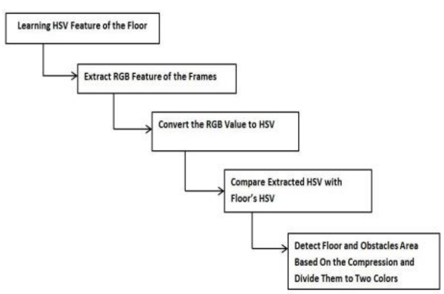

Figure 5: OpenCV algorithm

RoboRealm algorithm

RoboRealm is an application for use in computer vision, image analysis, and robotic vision systems. RoboRealm provides a Windows based GUI for experimenting with different modules that can be assembled in custom ways to achieve a desired result. It has compiled many image processing functions into an easy to use windows based application that we can use with a webcam. [12]

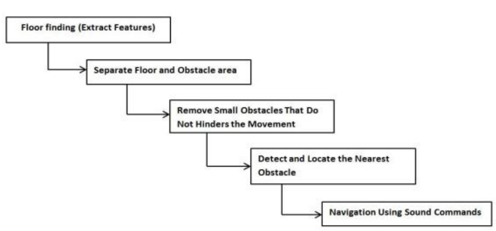

Figure 6: RoboRealm algorithm

VI. Implementation

OpenCV is written in portable C/C++. It should compile and run on any device that has a C/C++ compiler, the main issue is that OpenCV uses floating point math, while RoboRealm modules were written using many languages C++, C-Sharp, Java, Python, Visual basic and Visual basic.Net, to use the RoboRealm you need a Microsoft operating system.[13] Both OpenCV and RoboRealm allow developer to extend their libraries, in OpenCV you can write your own library in C/C++ and include it in OpenCV, in RoboRealm the Plugins allow you to create custom filters/algorithms that can be added into the RoboRealm image processing pipeline, you can simply write a processing code in another language and platform then provide it to RoboRealm using sockets (server hostname and port is needed, usually local host).

OpenCV Implementation: Unlike RGB, HSV separates "luma", or the image intensity, from "chroma" or the color information. This is very useful in our application since we want to deal only with the intensity component, and leave the color components alone.

Conversion from RGB to HSV could be done using equation.1.

$$H = \arccos \frac{\frac{1}{2}(2R-G-B)}{\sqrt{(R-G)^2-(R-B)(G-B)}} \qquad 1$$

In HSV the saturation and value (the color summarizes uses this model) are calculated using the following equations.

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \qquad .2$$

$$V = \max(R,G,B) \qquad .3$$

The code for converting between RGB and HSV is widely available and can also be easily implemented. In OpenCV, we write a code to capture frames from camera then convert them to binary color frames, where white area indicates where the ground is in each frame, black area indicates possible obstacles.

RoboRealm Software Implementation: RoboRealm is an application for use in computer vision, image analysis, and robotic vision systems. Using an easy point and click interface RoboRealm simplifies vision programming. With an inexpensive USB webcam and the PC vision projects can be implement. Image and/or video processing can be technically difficult. Home robots are continuously moving towards PC based systems (laptop, netbook, embedded, etc.) that have the power to support complex image processing functions. RoboRealm provides the software needed to get such a system up and running. RoboRealm has compiled many image processing functions into an easy to use windows based application that use can use with a webcam, TV Tuner, IP Camera, etc. RoboRealm software implementation done using modules, modules technique separate software functions into independent modules, such that each contains everything necessary to execute only one aspect of the processing functionality Figure 7 explain the relation between real area and camera scene. The real area

DOI : 03.IJIS.2015.1.4

scene dimension will be: depth 360 cm and a width of 210 cm, which equals 480 pixels and 640 pixels respectively. We choose rectangular shape with 110 pixel widths and 25 pixel height, these pixels represent around 36 cm width and 45 cm height in real area to sample floor area.

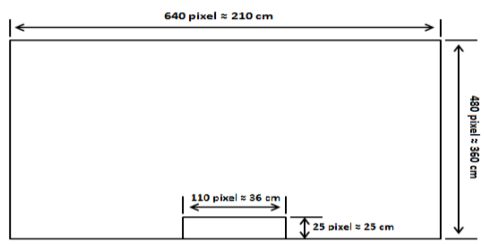

Figure 7: Real area scene view in cm with corresponding pixel location in the camera scene

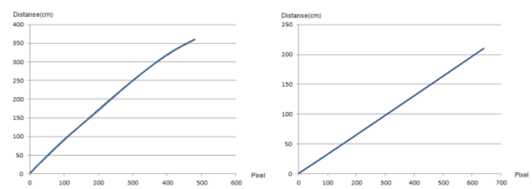

Figure 8: (a) Distance and vertical pixel location curve, (b) Distance and horizontal pixel location curve

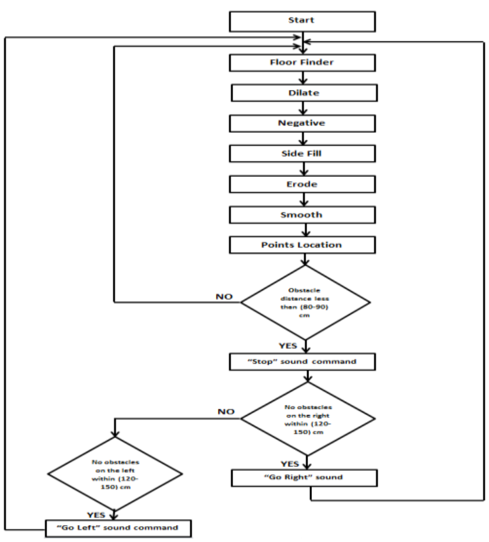

Figure 9: RoboRealm Flow chart

Conclusion and future work

Our system can detect obstacles in indoor environment and gives an acceptable results to guide visually-impaired people, this RoboRealm system could do detection even when ground is not consistent based on RGB/texture algorithms, when lighting is somehow changing the system won't fail which is counted as a very good point.

For future work, we can improve our system first in algorithms used in detection, we may develop more complex algorithms to better detect floor and obstacles in extreme cases like sharp lighting, another improvement will be the dealing with stairs problem because in testing we mentioned that our system cannot detect stairs, deploying the system on microprocessor also can be counted as a future work to help user to easily carry the system everywhere.

References

[1] Sylvie Treuillet, Eric Royer, Thierry Chateau, Michel Dhome, Jean-Marc Lavest "Body mounted vision system for visually impaired outdoor and indoor way finding assistance" CVHI 2007, M.A. Hersh (ed.).

[2] Ljupko Šimunović "Blind people guidance system" Faculty of Transport and Traffic Sciences University of Zagreb.

[3] Coombs, D., Roberts, K., "Bee-Bot: Using Peripheral Optical Flow to Avoid

[4] Obstacles" Intelligent Robots and Computer Vision XI, , SPIE Vol 1825 1992.

[5] Argyros, A., A., and Bergholm, F., "Combining Central and Peripheral Vision for Reactive Robot Navigation" Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages II:646-651, IEEE Computer Society Press,1999.

[6] Takeno, J., Shinogi, Y., Nishiyama, S., Mizuguchi, N., Sorimat, K., "Realization of a 3D Vision Mobile Robot that can Avoid Collision with Moving Obstacles," IEEE International Conference on Robotics and Automation, Sacramento, California April 1991.

[7] Langer, D., and Jochem, T., "Fusing Radar and Vision for Detecting, Classifying and Avoiding Roadway Obstacles", IEEE Intelligent Vehicles Symposium, 1996 pp333-338.

[8] Nelson, R. C., Aloimonos, J., "Obstacle Avoidance Using Flow Field Divergence", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol 11 No 10 October 1989. Syed Irtiza Ali Shah "Vision based 3D obstacle detection using a single camera" Georgia Institute of Technology August 2009.

[9] Jeffrey Ryan Lalonde monocular obstacle detection for moving vehicles University of Ottawa.

[10] Syed Irtiza Ali Shah "Vision based 3D obstacle detection using a single camera" Georgia Institute of Technology August 2009.

[11] Iwan Ulrich and Illah Nourbakhsh. Appearance-based obstacle detection with monocular color vision. In Proceedings of AAAI 2000, 2000

[12] http://www.roborealm.com. (Retrieved 05-01-2014)

DOI : 03.IJIS.2015.1.4