World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:7, No:11, 2013

Development of a Computer Vision System for the Blind and Visually Impaired Person

Rodrigo C. Belleza, Jr., Roselyn A. Maaño, Karl Patrick E. Camota, Darwin Kim Q. Bulawan

***Abstract*—**Eyes are an essential and conspicuous organ of the human body. Human eyes are outward and inward portals of the body that allows to see the outside world and provides glimpses into ones inner thoughts and feelings. Inevitable blindness and visual impairments may results from eye-related disease, trauma, or congenital or degenerative conditions that cannot be corrected by conventional means. The study emphasizes innovative tools that will serve as an aid to the blind and visually impaired (VI) individuals. The researchers fabricated a prototype that utilizes the Microsoft Kinect for Windows and Arduino microcontroller board. The prototype facilitates advanced gesture recognition, voice recognition, obstacle detection and indoor environment navigation. Open Computer Vision (OpenCV) performs image analysis, and gesture tracking to transform Kinect data to the desired output. A computer vision technology device provides greater accessibility for those with vision impairments.

***Keywords*—**Algorithms, Blind, Computer Vision, Embedded Systems, Image Analysis.

I. Introduction

BLINDNESS and visual impairment confines the ability to move around and control oneself in the environment. Individuals lacking optical perception may be due to physiological or neurological factors. The disability significantly interferes with one's ability to function independently and to perform activities safely and easily. Inexorable navigation hindrances become a real challenge most especially for persons with eye-related disabilities.

Canes and guide dogs were conventional means of way-finding for the blind and visually impaired. An article [1] describes the common method for navigating of visionless person using a white cane or walking cane. The walking cane is a simple and purely an object dedicated to detect static obstacles on the ground, uneven surfaces, holes and steps. This device is light, portable, but range is limited to its own size and not suitable for dynamic obstacles.

Based on the interdependence between the disabled owner and a dog, the training and the relationship to the animal is another way for the blinds to walk around. The dog is able to detect and analyze complex situations like stairs, detect potential danger and know paths. The blind is able to feel the attitude of his dog, analyze the situation and also give orders. But guide dogs are still far from being affordable, and their average living is seven years.

Rodrigo C. Belleza, Jr. is currently the Dean of College of Computer Studies, Manuel S. Enverga University Foundation, Lucena City, Quezon, Philippines (e-mail: rcbellezajr@gmail.com).

Roselyn A. Maaño is a faculty of College of Computer Studies, Manuel S. Enverga University Foundation, Lucena City, Quezon, Philippines (e-mail: roselyn_maano@yahoo.com.ph).

Karl Patrick E. Camota and Darwin Kim Q. Bulawanare BS Computer Science students of College of Computer Studies, Manuel S. Enverga University Foundation, Lucena City, Quezon, Philippines (e-mail: camota.kp@gmail.com, dkam31@yahoo.com).

The ideal correction of blindness and visual impairment may never be attained which consequence in an-over-or under-correction of vision resulted from expensive eye surgeries and transplants. However, according to an article [2], people who have been blind from birth make use of the visual parts of their brain to refine their sensation of sound and touch, according to an international team of researchers led by neuroscientists at Georgetown University Medical Center (GUMC).

Computer vision is a broad and complex field of study that touches the abilities of human eye's perceiving and understanding an image by electronic means. The researchers selected computer vision as an area of research to address the needs of the blind and visually impaired individuals. Various hardware and software components made up the system.

II. Objectives and Scope of the Study

The main objective of the study is to develop a prototype that will serve as an assistance to the blind and visually impaired utilizing the Kinect technology. The study emphasizes Microsoft Kinect device for developers that will be the eye of the system and Arduino microcontroller as a complementary part of the system.

Microsoft Kinect for Windows is a physical device that, when used with a computer and the software developer kit, provides companies and developers with the foundation they need to create interactive applications that recognize peoples' natural movements, gestures, and voice commands [3]. Kinect depth sensors are limited to 0.80m to 4.0m in the normal mode and 0.40m to 3.0m in the adjacent mode. The Kinect camera captured all movements within the ray of range and the frame rate yields up to 30 frames per second (fps). The camera sees objects in the anterior limited to 43° vertical by 57° horizontal field of view (fov) which was utilize in the indoor way finding of the system. The system is limited to a flat surface. In the obstacle detection, only ground and solid plane objects are detectable. The system cannot be used in direction-finding that requires steps or ladders. The system cannot accommodate a combination and series of disabilities of the deaf-mute-blind people.

III. Significance of the Study

According to an article published in Philippine Star [4], vision is a gift that enables everyone to see and appreciate the world's magnificence. The gift of vision takes everyone in the marvel of the rising sun and constellations in the night sky. Unfortunately, around 285 million people are visually impaired worldwide and 39 million of them are completely blind. In the Philippines, there is an estimated half a million blind people whose visual impairments are caused by different eye conditions and diseases.

The importance of the study focuses on developing a reliable and alternative solution to cater the needs of the blind and visually impaired individuals with the use of emerging technology of computer vision. Stakeholders of the system include blind and visually impaired individuals, foundations and institutions for the blind and engineering and computer science experts. The study will also serve as a reference for future researcher engrossed in the same field.

IV. Methodology

Research and development (R&D) are two intimately related processes by which new products and new forms of old products are brought into being through technological innovation [5]. Fig. 1 shows the research paradigm applied on the study which describes the integrated approach of conducting R&D research as described by Borg and Gall and Gall [6]. The graphical representation in Fig. 1 is composed of seven steps initializes from the theory building and research analysis of the system that needs to be developed up to the testing of the developed prototype.

The first step includes the development of new ideas and concepts, review of literatures, information gathering from books, journals and conference proceedings and construction of conceptual frameworks, methods to be used that is subjected to extensive analysis. Interviews with experts in the field were conducted to acquire the knowledge and data needed to construct the system. Personal communication via electronic mail with Mr. Vasily Tarasov aided the proponents in gathering the suitable hardware components and algorithm approaches in planning the creation of the project design in toeing the line with the objectives and feature of the system. Mr. Vasily Tarasov won Second Place in the Imagine Cup 2011 Software Design, with the project titled: Kinect for the Blind.

The data gathered in theory building became the ingredients in initializing the entire system. The prototype underwent designing, coding, testing, compiling and execution. The system has series of iterations that describes the development technique of prototyping. Prototyping lead the researchers to meet the specifics of the study. Prototyping is an experimental method which is intended to come up with a final prototype. The researchers identified the devices, hardware and software to be used in the system. The software and hardware components of the prototype describe the embedded nature of the developed system.

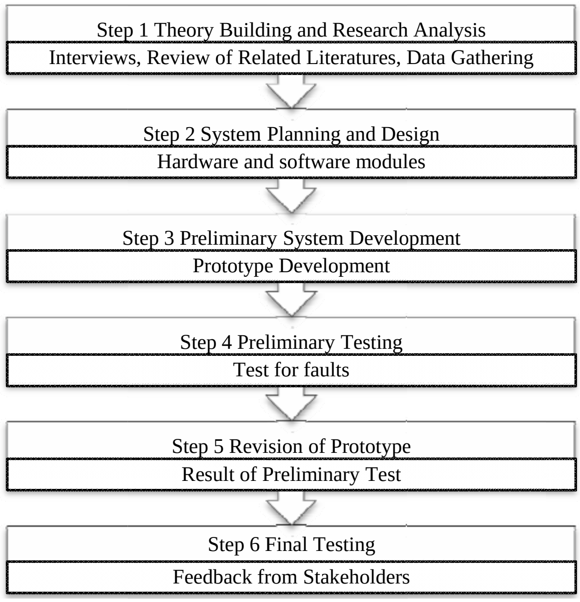

Fig. 1 R&D Methodology

V. System Functionalities

The system initializes in Microsoft Kinect Sensors which will process the data acquired from outside environment such as image and depth stream. All processed data will be optimized to produce a significant output which will be directed to the shaft less vibration motors for object detection and Microsoft Speech API for the recognition of markers using hamming codes.

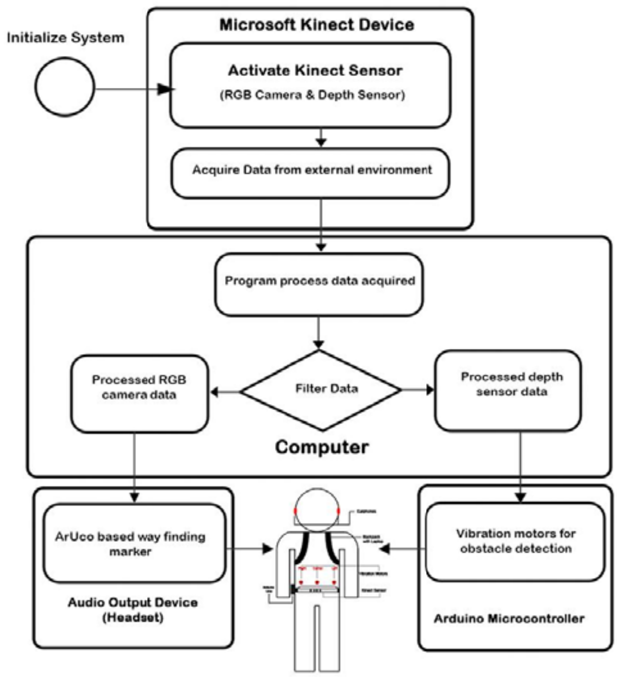

Fig. 2 Functionalities of the System

VI. Hardware Specification

A. Microsoft Kinect for Windows

The researcher maximized the features of the Microsoft Kinect for Windows utilizing Kinect Software Development Kit (SDK) and National User Interface (NUI). The Microsoft Kinect for Windows as in Fig. 3 contains an RGB camera that stores three channel data in a 1280x960 resolution. An infrared (IR) emitter produces infrared light beams and an IR depth sensor reads the infrared beams reflected back to the sensor. The reflected beams are converted into depth information measuring the distance between an object and the sensor. A multi-array microphone contains four microphones for capturing sound that made it possible to record audio as well as find the location of the sound source and the direction of the audio wave. A 3-axis accelerometer configured for a 2G range, where G is the acceleration due to gravity determine the alignment of the Kinect [7].

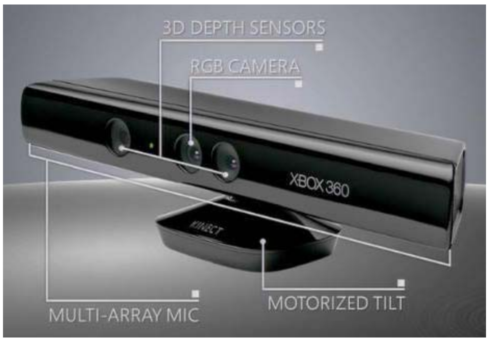

Fig. 3 Parts of Windows for Kinect

B. Arduino Microcontroller Board

Arduino is an open-source electronics prototyping platform based on flexible, easy-to-use hardware and software. The microcontroller on the board as mentioned in [8] is programmed using the Arduino programming language (based on Wiring) and the Arduino development environment (based on Processing).

C. Vibration Motor

Shaft less vibration motor by Precision Microdrives will serve as an indicator in the obstacle detection feature of the system. The motor's operating voltage ranges from 2.5V-3.8V. Three vibration motors held by patches are placed in front abdominal area of the user. The position of the vibration motors are determined by the position of the Kinect's camera. The schematic layout as in Fig. 4 illustrates the circuitry composed of Arduino Uno and vibration motor and discrete components such as 1K-ohm and 33-ohm resistors, 1N4001 diodes, 2N2222 transistors, and 0.1F ceramic capacitors.

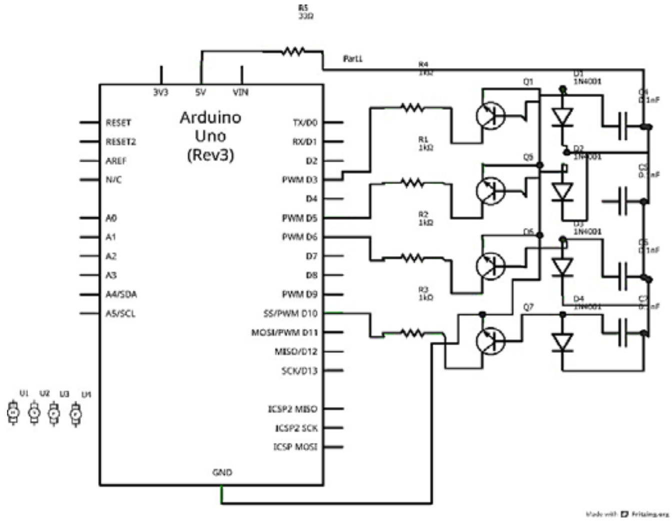

Fig. 4 Schematic diagram of Arduino Uno, vibration motor and discrete components

VII. Software Specifications

A. Kinect for Windows SDK 1.7

The researchers used the Kinect for Windows SDK 1.7 to access data such as depth and image stream from the Microsoft Kinect Sensor. The SDK provides tools and Application Programming Interface (APIs) for both native and managed applications. Microsoft Kinect is connected in Microsoft Visual Studio 2010 the system's development platform. In order to increase the speed of the image processing and manipulation, the Intel Thread Building Blocks (Intel TBB) take advantage of multi-core processor performance which offers a rich and complete approach to expressing parallelism in a C++ program.

B. OpenCV 2.4.6 and ArUco 1.2.4 Library

OpenCV known as an open source computer vision library developed by Intel will utilize all the programming functions and algorithms we used in the system. ArUco is a minimal C++ Library for detection of Augmented Reality markers based on OpenCV exclusively. OpenCV performs image analysis, and gesture tracking to transform Kinect data to the desired output. ArUco Augmented Reality Library was chosen to create and detect Hamming code-dependent markers for indoor navigation.

C. Image Analysis Algorithm

Various algorithms are carried out during the research such as Alexander Telea's Fast Marching Method, Depth Color Mapping, Multiple Threading, Bilinear Interpolation and Marker Detection and Recognition Process.

VIII. Testing

In the testing phase, the researchers tested the prototype among eight blind and visually impaired individuals. The prototype illustration in Fig. 5 includes Kinect sensor, waist assembly to mount the Kinect, laptop for processing the data, laptop backpack to hold the laptop and headphone for giving the user directions and Arduino Uno microcontroller board where vibration motors are mounted.

The blind or VI individuals moved around in a predefined environment. The blind and VI individuals relies on the system to identify which way to go. For the purpose of testing the obstacle avoidance, obstacles were set such as people walking towards the blind and VI and doors opening or closing. When an obstacle was determined by the system, the motors vibrate left, right or center. Left motor vibrates when it sensed obstacles on the left, right motor when obstacles are in the right and center motor when obstacle is in front of the blind. The indoor environment navigation was tested through ArUco-based markers set on the walls of the location. The blind were then ask to move around. The system can read the markers and use the text to speech method to tell the blind to turn right, left or where the blind is going through. The markers have equivalent codes embedded in the system. The assessment was furthered in a dim environment. The major disadvantage of dim light in the obstacle detection is what the system gets in terms of accuracy. The system was able to address the lighting condition and able to perform the same way with enough lighting situation.

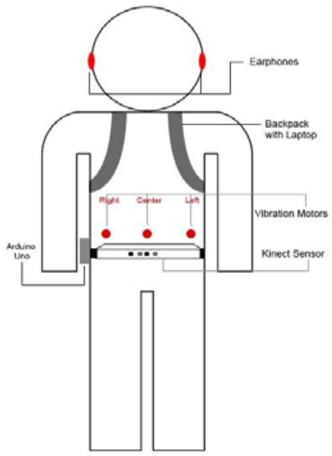

Fig. 5 Prototypesetup on blind and VI individuals

IX. Results

A. Depth Color Mapping

The depth frame is consists of 11-bit depth information in one unsigned integer per pixel and set the depth stream resolution in 640x480 pixels. Each depth frame state array has a total of 307,200 bytes that values each of 11-bit in size. The depth information is limited to 0 - 2047 and a plain meaningless black and white image in Fig. 6.

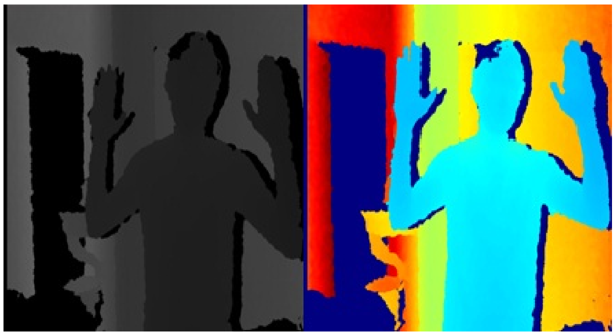

Fig. 6 Plain Depth Data (Left) and Processed Depth Data with RGB Values (Right)

B. Multiple Threads for Image Processing

One of the main problems in real-time applications is the system average speed performance where it takes few seconds to update a new frame due to large image data being processed by more than one high level image processing algorithm. The Kinect depth and image stream typically runs on 30fps on a regular basis without image processing. But when image processing occurs, both depth and image stream slows down from 30fps down to 10-15fps. To solve this problem, the researchers managed to use multiple threads and allocating the separate image data on a separate memory buffer.

C. Bilinear Interpolation

In order to remove the Kinect shadows, the researchers used the Fast Marching Method (FMM), a popular in painting algorithm by Alexander Telea, the result is similar to Fig. 7. The Kinect RGB Camera, IR camera, and IR projector are positioned on different parts of the device, so when one aligns the depth information with the point of view of the camera, the dark areas are either out of range from the point of view of the IR camera, or are absorbing IR light resulting to no depth information [10].

Fig. 7 Before (left) and after (right) of the in painting Method

D. Accessing Pixel Intensity Values

The 8x4 matrix was converted to a single channel grey scale image in order to get the pixel intensity value. Each pixel contains a value from 0 to 255 which indicates that 0 is equal to black and white is equal to 255.

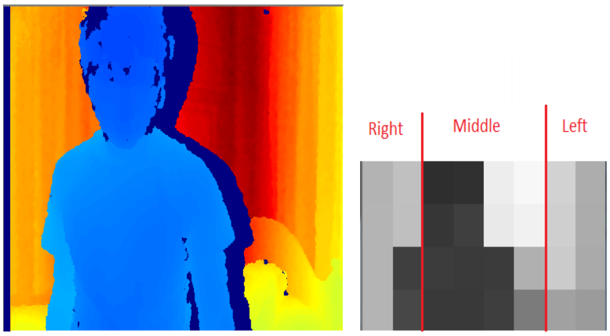

Fig. 8 Inprinted depth image frame

TABLE I
INTERPRETED VALUES OF THE SCALED 8X4 MATRIX

179	192	46	48	237	247	210	173
180	191	54	63	233	241	207	172
180	63	60	58	59	176	203	170
180	71	57	57	62	123	161	155

The matrix in Table I corresponds to the numerical pixel value of shaft less vibration motor during execution. If obstacle was detected in a distance less than one meter, the corresponding pixel value is less than 60. The total intensity per region from right to left, sends trigger commands through serial communication in the Arduino UNO with 3 vibration motors connected to different pin outputs. The higher the intensity values, the faster the speed of the vibration motors.

E. Marker Recognition Process

The ArUco Augmented Reality Library was utilized to detect markers for indoor navigation. The library relies on the use of coded markers. Each marker has a unique code indicated by the black and white colors. The library detects borders, and analyzes into the rectangular regions which are likely to be markers. Each marker has an internal code given by 5 words of 5 bits each. The code employed is a minor modification of the Hamming Code. Each word has only 2 bits of information out of the 5 bits employed. The other 3 are employed for error detection which resulted to 1024 different identifications [9].

X. FINDINGS AND CONCLUSION

The study presented a different way of using the Microsoft Kinect RGB-D Sensor for navigation which will assist the blind and visually impaired.

In general, the system is an alternative way to replace the white cane or even the walking dog to facilitate the navigation or walking of the blind person.

OpenCV is a distinct tool for computer vision which was designed for computational efficiency and with strong focus on real time applications and contributed in the development of the prototype. Kinect SDK and OpenCV use C++ API that enables the researchers to integrate the Kinect Sensor with the computer vision library for implementing different functions and algorithms.

Memory allocation and Intel Thread Building Blocks (Intel TBB) was used to enable multi-core processing that enables a faster rate measured in frames per second.

In order to remove the Kinect shadows, the researchers used the Fast Marching Method (FMM), a popular in painting algorithm by Alexander Telea.

Bilinear interpolation uses only the 4 nearest pixel values which are located in diagonal directions from a given pixel in order to find the appropriate color intensity values of that pixel. Bilinear interpolation considers the closest 2x2 neighborhood of known pixel values surrounding the unknown pixel's computed location and takes a weighted average of these 4 pixels to arrive at its final, interpolated value. The weight on each of the 4 pixel values is based on the computed pixel's distance in 2D space from each of the known points [11].

ArUco applies Adaptive Thresholding to detect borders and analyze borders into the rectangular regions. The adaptive thresholding process aims to filter out unwanted borders.

REFERENCES

[1] Shah, C. Bouzit, M., Youssef, M. &Vasquez, L. Evaluation of RU-Netra - Tactile feedback navigation system for the visually impaired. Page(s):72 – 77; ISBN:1-4244-0280-8; INSPEC Accession Number: 9273838; DOI: 10.1109/IWVR.2006.1707530.
[2] Science Daily Georgetown University Medical Center. People blind from birth use visual brain area to improve other senses: Can hear and feel with greater acuity. Retrieved from http://www.sciencedaily.com, October 10, 2010.
[3] Product Features | Microsoft Kinect for Windows. (n.d.).Retrieved from http://www.microsoft.com.
[4] A second hard look at blindness. The Philippine Star. Retrieved from http://www.philstar.com. December 18, 2012.
[5] W. K. Holstein. Research and Development. Retrieved from http://global.britannica.com. 2013.
[6] Borg, Gall and Gall. Education Research: An Introduction. Massachusetts: USA, 2011.
[7] Parts of Windows for Kinect. Retrieved from http://msdn.microsoft.com.
[8] Arduino microcontroller. Arduino official page. http://arduino.cc, 2013.
[9] ArUco: a minimal library for Augmented Reality applications based on OpenCv. Retrieved from http://www.uco.es.
[10] Kinect depth inpainting and filtering. Retrieved from http://www.radfordparker.com.
[11] Surbhi Syal, Sandeep Sood , Sunny Sharma, Er. Navneet Randhawa. Segmented character recognition using neural networks. Retrieved from http://www.ijera.com. Vol. 1, Issue 4, pp.1731-1735ISSN: 2248-9622.