# Fuzzy-Rule-Based Object Identification Methodology for NAVI System

**R. Nagarajan**

*Artificial Intelligence Research Group, School of Engineering and Information Technology, Universiti Malaysia Sabah, Locked Bag 2073, Kota Kinabalu 88999, Sabah, Malaysia*
*Email: nagaraja@ums.edu.my*

**G. Sainarayanan**

*Artificial Intelligence Research Group, School of Engineering and Information Technology, Universiti Malaysia Sabah, Locked Bag 2073, Kota Kinabalu 88999, Sabah, Malaysia*
*Email: jgksai@ums.edu.my*

**Sazali Yaacob**

*School of Mechatronic Engineering, Northern Malaysia University College of Engineering, Kubang Gajah, Arau 02600, Perlis, Malaysia*
*Email: s.yaacob@kukum.edu.my*

**Rosalyn R. Porle**

*Artificial Intelligence Research Group, School of Engineering and Information Technology, Universiti Malaysia Sabah, Locked Bag 2073, Kota Kinabalu 88999, Sabah, Malaysia*
*Email: rlyn39@pasca.ums.edu.my*

We present an object identification methodology applied in a navigation assistance for visually impaired (NAVI) system. The NAVI has a single board processing system (SBPS), a digital video camera mounted headgear, and a pair of stereo earphones. The captured image from the camera is processed by the SBPS to generate a specially structured stereo sound suitable for vision impaired people in understanding the presence of objects/obstacles in front of them. The image processing stage is designed to identify the objects in the captured image. Edge detection and edge-linking procedures are applied in the processing of image. A concept of object preference is included in the image processing scheme and this concept is realized using a fuzzy-rule base. The blind users are trained with the stereo sound produced by NAVI for achieving a collision-free autonomous navigation.

**Keywords and phrases:** blind navigation, edge detection, object identification, fuzzy-rule-based, stereo sound.

## 1. INTRODUCTION

The World Health Organization estimated that around 180 million people worldwide are visually disabled. Of those, between 40 and 45 million people are blind [1]. Blind people's navigation is restricted and sometime hazardous because they do not receive enough information on the objects or obstacles in their environment. Electronic travel aids (ETAs) are electronic devices designed to aid the navigation of blind people. ETAs design depends highly on the type of sensor used in the system, the method of conveying information to the blind, and also the system hardware. Most of the early ETAs use ultrasonic, infrared, or laser sensors for obstacle detection [2]. Since 1991, the development of vision substitution system using optical devices has become eminent.

*The vOICe*, Vuphonics and Optophone are the only most important earlier ETAs that use optical sensors as input devices. The vOICe was developed by Meijer in 1991 [3]. The working concept of The vOICe is based on "image-to-sound" conversion wherein the system captures images in front of the blind user and the image properties are transformed into sound patterns. The image is resized to $64 \times 64$ pixels and grey-scaled into 16 levels. The sound pattern is generated using sine wave; the loudness depends on the intensity value of each pixel in the image. Another similar invention is the Vuphonics, which is developed by Dewhurst [4]. This system attempts to convey the feature of visual image through coded sound effect and their tactile equivalent. Efforts are made to realize features such as the colour, texture, distance, as well

as the entities from the captured image. The Optophone developed by Picton and Capp is another device that attempts to map image to sound [5].

Navigation assistance for visually impaired (NAVI) is a vision substitution system that is also based on image-to-sound concept. This system has been developed to identify objects or obstacles in indoor environment such as inside the building and along a corridor. One of the main differences between NAVI and the other vision sensor-based aids is that the implementation of image processing methodology in NAVI facilitates more importance to the properties of objects rather than the background. The captured image from the vision sensor was resized to $32 \times 32$ pixels and quantized into four grey levels. Objects were identified with grey values. The identified objects were then enhanced and the background was suppressed using a clustering algorithm [6]. The processed image was mapped into stereo sound patterns; the amplitude of the sound was a function on the grey intensity of the image. The NAVI system was tested on blind volunteers to get the response as well as the suggestions regarding the pleasantness and logical meaning of the sound produced.

The objective of this paper is to improve the image processing methodology in the NAVI system. In the earlier image processing, the objects in the captured image were identified with their grey values. Due to lighting effects, a single object may have more than one grey level. Hence, there were possibilities that objects in the image were not highlighted compared to background. Noise elimination procedure was not undertaken in the earlier version. Furthermore, the grey-level quantitization in the processing would result in a reduction of information in the image. This information is inevitable for object identification.

This paper highlights how these drawbacks have been eliminated in the new design of image processing methodology. In the proposed procedure, a noise elimination stage is also incorporated. A concept of object preference is included so that the blind can identify the presence of nearest object for a collision-free navigation.

## 2. HARDWARE OF NAVI

NAVI system developed for vision substitution has a headgear mounted with the vision sensor, a pair of stereo earphones, a single board processing system (SBPS), and a specially designed vest for housing the SBPS. The SPBS selected for this system is a PCM-9550F version with an Embedded Intel low-power Pentium MMX 266 MHz processor, 128 MB SDRAM, 2.5-inch lightweight hard disk, two universal serial buses, and an RTL 8139 sound device chipset, all assembled in a microbox PC-300 chassis. The weight of SBPS is 0.7 kg. Constant 5 V and 12 V supply for SBPS are provided from the batteries placed in the front packets of the vest. The weight of battery is approximately 0.2 kg. Vision sensor selected for this application is a digital video camera, Kodak DVC325 [6]. Figure 1 shows a blind individual carrying the headgear and the processing system in the vest.
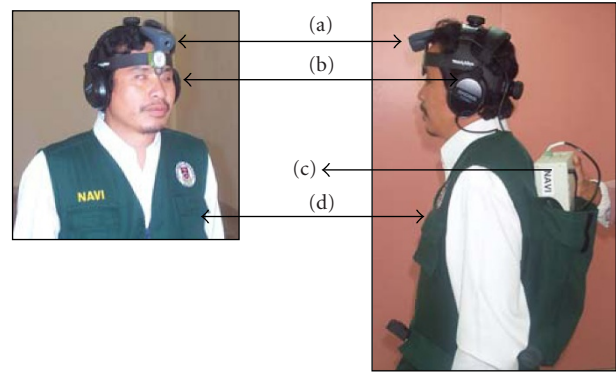


FIGURE 1: Blind volunteer with NAVI system. (a) Digital video camera fixed in headgear. (b) Stereo earphones. (c) Single board processing system with chassis. (d) NAVI vest.

## 3. IMAGE PROCESSING METHODOLOGY

The image captured by the digital vision sensor has to be processed in real time and transformed into an acoustic pattern to the earphones of the blind user. Since the processing is done in real time, the time factor consideration is critical. Real time, we mean here, is the continuous sampling and processing of the images of the scene and converting as sound patterns to the earphones at a specific rate. In the developed prototype equipment, a rate of 1 (or 2) image frame per second is employed. The rate of sampling depends on the duration of image acquisition, image transfer to processing system, image processing, sound generation, sonic transfer to earphones, and the duration of acoustic information heard through earphones for each image. However, the sampling rate mostly depends on time duration of acoustic information heard though earphones for each image. Initially, during training phase of the blind, the duration of acoustic information is set longer for each image. As the blind becomes familiar with the sound generated from the equipment, the duration of the acoustic information for each image and hence the time gap between each sampling can be reduced. At the training stage, the sampling time is fixed to approximately 1 second. Then the computational time of the processing module has to be much less than 1 second so that the blind has sufficient time to hear and understand the sound [6].

Most of the earlier works on image-to-sound-based system for the blind in the literature do not consider any suitable image processing procedures. In these works, information from the unprocessed image is fed to the auditory system of blind person. This excessive information causes a difficulty to the blind people in understanding the feature of objects in the image. In this proposed work of NAVI, the amount of information to be delivered to the auditory system is reduced so that only the required information regarding the features of objects is given to the blind.

In the case of an industrial vision, the object location and its features can be easily identified since the object models lighting, and other parameters are known a priori. In vision substitution system such data cannot be known a priori due to highly changing environment in front of the blind.

In Figure 2, the block diagram of the proposed image processing methodology is shown. The captured colour image is resized to $32 \times 32$ [6]. The reason for selecting $32 \times 32$ size is mainly to achieve the smallest real-time image processing without compromising the quality of objects in the image plane. This selection of resizing is also important in restricting the amount of information that human ear can assimilate per second. Canny edge detector is used to detect edge locations in three-colour components in RGB image and the results are combined to give overall edges of the image. A method of edge-linking is proposed to combine the broken edges and the edges at the boundary of images. Unnecessary edges are removed using a noise removal procedure. Region inside the closed edges are considered as objects. Fuzzy rules are applied to assign a preference to the identified objects. Each of these processing steps is discussed in detail in the following sections.

### 3.1. Edge detection

Edge detection is one of the most important human vision properties. Often human identifies the object with its boundaries and shape. The goal of edge extraction is to provide useful structural information about object boundaries. From the edges, the object properties such as area, perimeter, and shape can be measured for object identification [7]. Laplacian, Roberts, Sobel, and Prewitt are some of the edge detector operators [8]. These operators can detect edges by using convolution masks that represent the "ideal" edge step in various directions. The main disadvantages of these operators are their scale dependence and noise sensitivity. Canny edge detector is one of the optimum edge detecting methods for step edges corrupted by white noise [9]. Canny edge detector has better performance in edge detection than the other methods. In the earlier image processing methodology of NAVI system, the boundary of objects is detected using a neural network-based Canny edge operator. The size of the filter, threshold, and the ratio of threshold are derived by applying a feed-forward neural network. In some cases, the edge detectors applied to greyscale image do not provide enough information in identifying a meaningful object. Many researchers have extended their attention to colour image processing since more edge information can be extracted from colour images [10, 11].

In this work, an attempt is made to use colour information for edge detection in NAVI system. Generally the edge detection method on colour images can be classified into three categories. First, the edges are detected in each colour channel (e.g., R, G, and B components), respectively. The results are then combined. Second, edges are detected in the luminance channel only. Third, edges are detected based on the gradient in the three-channel image field [12]. More recently, Wesolkowski et al. examined a comparison of colour edge detection performance in several colour spaces [13]. In this paper, the edge detection is employed in RGB colour space. The colour image is separated into red component, green component, and blue component. Canny edge detector is used to determine the edge location of each colour component. Assuming all edges present in each colour component as
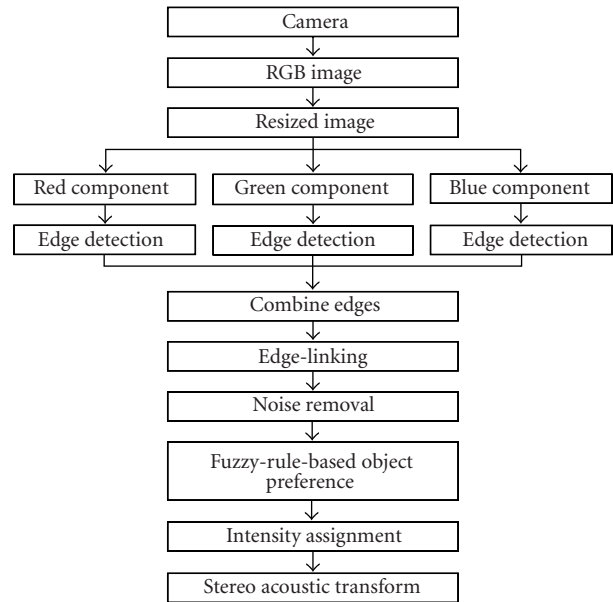


Figure 2: Block diagram of the image processing methodology.

desired edges, they are combined using an "OR" operator. The implemented Canny edge detector has been found adequate for NAVI application. However, it is realized that edge detection alone is not enough to extract the object from the image. Further processing is undertaken to identify and connect the broken edges to form meaningful objects.

### 3.2. Edge-linking

Images of real scene naturally contain objects, which are incomplete or distorted by various factors such as the lighting effects and the positioning of the object. Due to these factors, edges in the real images seldom form the closed and connected boundaries that are required for the object extraction. An edge-linking process is required to assemble these edges into meaningful edges. Based on the analysis performed in several experiments, the combined edges produce gaps of less than three pixels in the image. As a solution, an edge-linking algorithm is undertaken to connect broken edges with two gaps. The combined edge has to be thinned for preprocessing. This is to provide an edge width of one pixel. Each edge is then analyzed in $3 \times 5$ neighbourhoods for horizontal direction and $5 \times 3$ neighbourhoods for vertical direction. This candidate edge must smoothly connect the broken edges. Figure 3 illustrates the candidate edges for horizontal and vertical directions.

In the experimentation, it is also noted that the edge detector fails to track edges at the border of image. In vision substitution system, all objects regardless of their positions must be considered. Some of the objects may be located at the border of the image and so it is not in complete form. Thus the next approach in edge-linking aims to connect the broken edges at the border of image. Most of the important edges will be missing in the top and bottom row as well as in the first and last column of the image. The edge-linking is
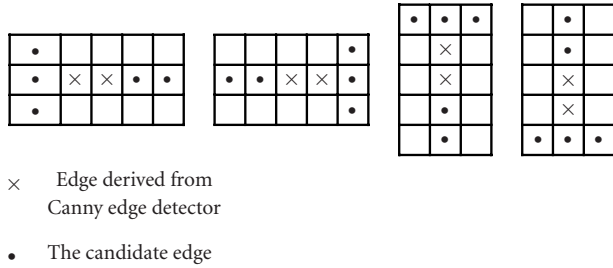
FIGURE 3: Candidate edges for horizontal and vertical directions.

× Edge derived from Canny edge detector

• The candidate edge



Outside iris area region

Iris area region

FIGURE 4: The "iris area" and the "outside iris area" of the image.

done for each border to form edges at this border. For the top border, edges in the second row of the image are scanned. If broken edges are found, pixels above the scanned edges are set as foreground. The same method applies for other three borders of the image. After the necessary edge is formed at the border of the image, edges present at the border are connected so that a complete object boundary is formed. All edges in the image are labelled to ensure that only edges with the same label will be connected. Once the edges are labelled, each edge in the image border matrix is scanned. Starting at the first scanned edge, the next edge is located and identified as the end edge if it has the same label with the first edge. Each pixel between the two scanned edges is presumed to belong to a straight edge segment and this pixel is set to foreground. If there are more than two edges with the same label present at the border, the process is omitted.

### 3.3. Noise removal

The goal of noise removal is to remove the extraneous edges present in the image without affecting the desired objects. Some basic morphological operations are used in this stage. Morphology is a technique for extracting image components, which are useful for region shape representation. In the morphological operation, a pixel that has value 0 (*off* pixel) is considered as background and a pixel that has value other than 0 (*on* pixel) is considered as foreground. Object is located as a set of on pixels in an image that form a connected group [14, 15]. Region inside the closed edge is identified as an object. Therefore, this region has to be set to foreground. Dilation and erosion operations are undertaken to smooth the object images. A disk structuring element with the size of one pixel is created. This structuring element is used in erosion operation to remove one pixel from around the boundary of all objects. As a result, the extraneous edges present in the image will be eliminated and the objects will shrink. To restore the objects to their original size, dilation operation is applied to the eroded objects using the same structuring element.

### 3.4. Fuzzy-rule-based object preference

The fundamental task involves in this stage the determination of the object of interest in the enhanced image. Usually in the industrial machine vision, objects of interest are found by eval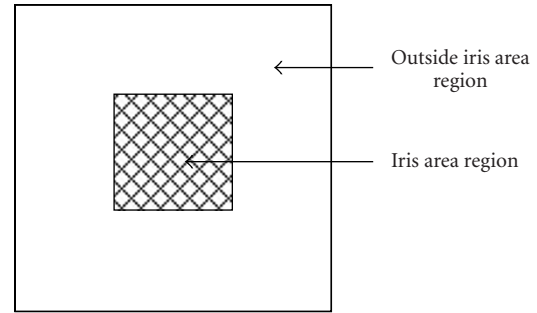uating certain object features such as size, colour, texture, and position. The object of interest and its features information are known and specified before the recognition task can be performed [16]. On the other hand, the selection of object of interest for a blind navigation application is uncertain and time varying. This is due to the constant shifting of the headgear-mounted camera orientation by the blind people. Significant difficulties arise as the features of objects in their environment keep on changing as the blind person moves and thus it is not possible to identify the object of interest by evaluating the fixed object feature. To resolve these uncertainties, certain properties are taken into consideration. One of the considerations is the creation of a new parameter for objects in the enhanced image. The parameter chosen here is the preference of the object. The object preference acts as a guidance to blind people to determine object of interest in the image. Moreover, by having different preferences, the blind user can easily discriminate the object properties in the environment.

To simplify the recognition task as well as to avoid confusion to the blind user, the preference levels to be assigned to objects in the image should be in small number. In this work, three preferences are set: the high preference, the medium preference, and the low preference. The high preference level indicates that the object is highly preferred and it is the object of interest in the image. Object with the medium or small preference level indicates that the object is less preferred than the object of interest but still has importance for navigation purposes. To compute the preferences of object, fuzzy-rule-based system is employed. Since many of the basic concepts in recognition task are ambiguous in nature and cannot be defined precisely, the fuzzy-rule-based system appears to be a good choice to deal with such uncertainty issues.

The designed fuzzy-rule-based system has three input features namely object size $I_1$, pixel distribution outside iris area $I_2$, and pixel distribution inside iris area $I_3$. The object size is the total pixels within the object in the image. Since the size of image is $32 \times 32$ pixels, the maximum total pixel that an object can take is 1024. In this paper, the term "iris area" is used to indicate the central $(8 \times 8)$-pixel area of the image. The pixel distribution outside the iris area is defined as the total object pixels which are not located at the central area of the image. The pixel distribution inside iris area is defined as the total object pixels that are located at the central

area of the image. Figure 4 illustrates the image used in this paper and the region of iris area and the region outside iris area. Each input feature is expressed using three membership functions namely small, medium, and big. The membership functions such as small and big are expressed using trapezoidal curve. For medium membership, a triangular curve is used. The output, object preference O, has three triangular curve memberships. Table 1 shows the inputs and the output of the fuzzy-rule-based system. The defuzzification is performed using centroid method.

A total of 27 rules are derived for this system. The rules are based on certain observations such as human visual consideration and also from the experience with the blind person. Since NAVI system is to be used by human blind, it is necessary to consider human preference regarding the selection of object of interest in the image. The first property is that if human concentrates on a particular object, other objects that surround the object of interest has less preference. The background goes a little out of concentration and hence is less focused. If there is only one object in the scene, the object would be detected as high-preference object regardless of the size and the position of that object. The objects can be detected before the object features are extracted. For this condition also, it is not important to evaluate the attributes of object features. Thus this condition can be directly evaluated without using the fuzzy-rule-based system.

The second property is based on the fact that the image in the real world usually contains more than one object. For this property, the fuzzy-rule-based system is essential to evaluate the attributes of object features. To provide a collision-free navigation system, the size and the location of object are considered as important features. In this case, a bigger object is more important than a smaller object. Apart from that, the object which is located at the centre of sight is considered more important than objects which are located away from the centre of sight. Another important note is that as the blind person gets nearer to any object, the size of object gradually becomes larger. All the above discussed conditions are taken into consideration while developing the fuzzy rules. Some of the fuzzy rules are given as follows.

(i) Rule 1. If $I_1$ is small, $I_2$ is small, and $I_3$ is small, then O is low.
(ii) Rule 3. If $I_1$ is small, $I_2$ is small, and $I_3$ is big, then O is high.
(iii) Rule 10. If $I_1$ is medium, $I_2$ is small, and $I_3$ is small, then O is low.
(iv) Rule 11. If $I_1$ is medium, $I_2$ is small, and $I_3$ is medium, then O is medium.
(v) Rule 19. If $I_1$ is big, $I_2$ is small, and $I_3$ is small, then O is medium.
(vi) Rule 21. If $I_1$ is big, $I_2$ is small, and $I_3$ is big, then O is high.

### 3.5. Intensity assignment

The result of the fuzzy-rule-based system produces three outputs which are low, medium, and high preferences. These

TABLE 1: Inputs and output of fuzzy-rule-based object preference.

|  | Variable | Membership function |
|---|---|---|
| Input | Object size $I_1$ | Small |
|  |  | Medium |
|  |  | Big |
|  | Pixel distribution outside iris area $I_2$ | Small |
|  |  | Medium |
|  |  | Big |
|  | Pixel distribution inside iris area $I_3$ | Small |
|  |  | Medium |
|  |  | Big |
| Output | Object preferences O | Low |
|  |  | Medium |
|  |  | High |

outputs are referred for assigning the objects with different grey intensities. Dark grey intensity is assigned for a low-preference object. A medium-preference object is shaded into light grey intensity and a high preference into white intensity.

### 3.6. Testing

In experimentation, 200 colour images have been tested to evaluate the proposed methodology. The testing is done offline and online. The result of each proposed stage in image processing methodology is shown in Figure 5. Figures 5a and 5b show the results of proposed methodology for indoor images. In Figure 5a, four objects with different sizes are shown. The respective fuzzy output images are shown in the sixth column. From the output, one object is highlighted with high intensity and others with low intensity. Figure 5b shows another indoor image. In this figure, three objects can be detected and object in the centre of image has high preference. Other objects have low preferences. Figures 5c and 5d illustrate the results of proposed methodology using outdoor images. In Figure 5c, only one object is detected and its edges are shown in third column. Using the edge detection, the line on the floor is also detected. By eliminating the extra edges and enhancing the object region, the output image shows only the object and the background. Since only one object is detected, this object is highlighted with high-level intensity. In Figure 5d, there is one object in the centre of image and the other at the left side of the image. The edge image produces complex edges due to the colour pattern on the floor. However by eliminating edges and enhancing the object region, the shape and size of the objects can be inferred. Using the fuzzy rules, the object in the centre of image is given a high preference and the other a lesser preference.

## 4. STEREO ACOUSTIC TRANSFORMATION

The final processed image matrix $I_p$ has to be transformed into sound patterns. The frequency or pitch of the human
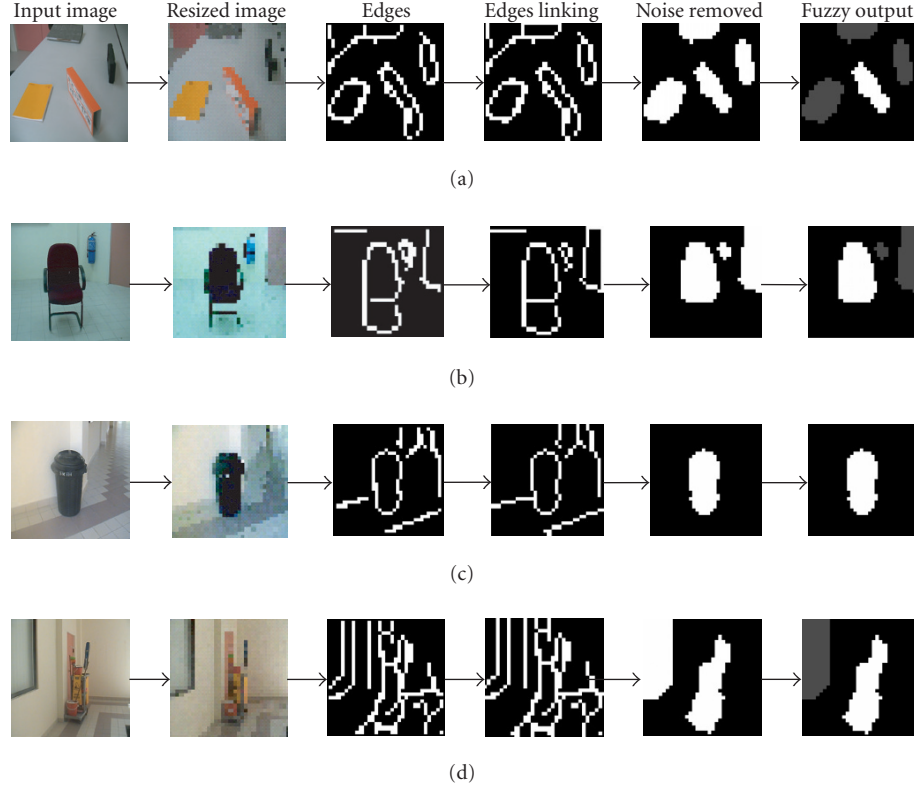
FIGURE 5: Results of the proposed methodology.

audible range is from 20 Hz–20 kHz. It is reported, however, that the human auditory is more sensitive in frequency range of 20 Hz–4000 Hz [17]. Since the human auditory system is more sensitive for discrimination in the low-frequency range rather than the high-frequency range, the frequency band is selected to be in the lower side of the frequency range. The vertical position of the image pixel is inversely related to pitch and the pixel intensity is converted into loudness of the sound. The frequency variations in the vertical position are designed to be audibly differentiable. The range of frequency for the sine wave sound generator is set at $F_L = 150$ Hz and $F_H = 4000$ Hz. This range and increment of frequency are selected by one of the blind volunteers so that he/she is comfortable with the repeated sound without the deterrence of the environmental sounds. Such specification seems to be fitting to other volunteers as well.

Let

(i) $f_0$ be the fundamental frequency of the sound generator,

(ii) $G$ a constant gain,

(iii) $F_D$ the frequency difference between the adjacent pixels in vertical direction.

The changes in frequency corresponding to $(i, j)$th pixel in the $32 \times 32$ image matrix is given by

$$F_D = f_i - f_{i-1}, \tag{1}$$

where "$i$" is the row number $= 1, 2, 3, \ldots, 32$. $f_i$ is the frequency of the sine wave for pixels in row "$i$" and is represented by

$$f_i = F_L + GF_L(32 - i). \tag{2}$$

In the proposed system, the frequency is linearly varied by maintaining $F_D$ as a constant.

The generated sound pattern is hence given by

$$S(j) = \sum_{i=1}^{32} I_P(i, j) \sin \omega(i)t; \quad j = 1, 2, \ldots, 16, \tag{3}$$

where

(i) $S(j)$ is the sound pattern for column $j$ of the image,

(ii) $t = 0, \ldots, D$, where $D$ depends on the total duration of the acoustic information for each column of the image,

(iii) $\omega(i) = 2\pi f_i$, where $f_i$ is the frequency corresponding to row, $i$.

The range $\{i = 1, 2, 3, \ldots, 32; j = 1, 2, \ldots, 16\}$ indicates one half of the image. The sound patterns created from the left half side of the image is given to the left earphone and sound patterns of the right half side to the right earphone simultaneously thus creating a stereo acoustic transformation. The scanning is performed from leftmost columns towards the centre and from rightmost columns towards the centre.
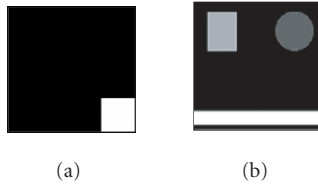
(a)          (b)

FIGURE 6: Examples of simulated images and the description of their sound patterns. (a) Low-frequency band of sound signals only in the right earphone; left earphone has no sound signal. (b) High-frequency band of sound signals for rectangle in the left earphone with higher amplitude and with a time lag from the start of sound. High-frequency band of sound signals for circle in the right earphone with lower amplitude. High-amplitude low-frequency signals for bottom rectangle in both the earphones for full duration.

## 5. TRAINING

The developed scheme of stereo acoustic transform was tested on a visually handicapped volunteer. The sound patterns from the real-time images are initially complex to understand and to categorize. For example, in learning Braille codes, the blind individuals have to undergo intensive and systematic courses. In a similar way, to start with the training of sound from the NAVI, it would be appropriate to first train the blind user with simulated images. The training procedure should create an interest and involvement in the blind volunteers. They should feel more and more confident as they put in more attention towards using the system. As a systematic procedure, the blind volunteer is first trained with the sound produced by the simulated images developed in Microsoft Windows paint. Initial training is started with the shapes like square, triangle, circle placed in different horizontal and vertical orientations in the image frame. The blind volunteer was asked to carefully listen to the sound. The logical meaning such as the shape of objects and their orientation in the image frame were explained to them in each training session. Figure 6 shows some simple simulated images with corresponding sound descriptions used in the initial stages of training. Figure 7 shows a set of test images with corresponding sound descriptions used for later stages of training.

After a good level of training, the blind volunteers were able to identify complex simulated images, the objects in indoor environment and the objects in simple outdoor environment. The volunteers were able to walk along the corridor with restricting obstacles. Slowly moving objects and their directions of motion were also inferred by the blind with the sound produced from the NAVI system. For understanding complex outdoor environments from routine life, the blind need continuous training and a long experience with the system. The efficiency of the blind in perceiving the objects is found to be increasing with continuous training. In this research, information regarding depth of the object is not considered, and it can be developed using stereo cameras. However in the single camera NAVI, by comparing the sound patterns from relative distances between the blind person and the object, information regarding distance of objects can be manipulated by the blind person with experience.
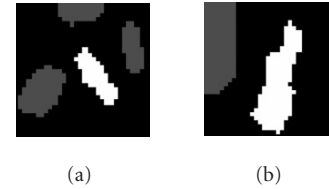


(a)          (b)

FIGURE 7: Examples of real-life images and the description of their sound patterns. (a) A band of medium-to-low frequency of sound with low amplitude from the left earphone, followed with a band of high-frequency sound with lower amplitude. A band of high-to-medium frequency sound with low amplitude from right earphone, followed with a band of medium-to-low frequency band with high amplitude. A band of high-frequency sound signal with low amplitude is inferred in the second half of time duration. (b) A band of high-to-medium frequency sound signals in the left earphone with lower amplitude in the first half of time duration. This is followed with a band of low-frequency sound signal with high amplitude until the end of time duration. A band of high-to-medium frequency with high amplitude sound signals in the right earphones in the second half of the time duration.

## 6. CONCLUSION

The main objective of this paper is to identify objects in a captured image. Objects are identified by their closed boundaries. Several approaches of object extraction involved in this paper include the edge detection, the edge-linking, and the noise removal. Once the objects are extracted, a fuzzy rule base is implemented for object identification. In human vision system, often the object of interest is located at the centre of the sight. The proposed fuzzy-rule-based methodology provides a preference to each object in the image. The processed image is sonified using a stereo sound procedure. The developed sonification procedure was tested on visually handicapped volunteers. With continuous training, the blind user can identify the location of objects through the sound pattern produced from the processed image.

## REFERENCES

[1] "World Sight Day: 10 October", World Health Organization, http://www.who.int/mediacentre/releases/pr79/en/print.html.

[2] F. Wong, R. Nagarajan, S. Yaacob, A. Chekima, and N.-E. Belkhamza, "Electronic travel aids for visually impaired—a guided tour," in Proc. Conference in Engineering in Sarawak, pp. 377–382, Sarawak, Malaysia, May 2000.

[3] P. B. L. Meijer, "An experimental system for auditory image representations," IEEE Trans. Biomed. Eng., vol. 39, no. 2, pp. 112–121, 1992.

[4] D. Dewhurst, "The Vuphonics home page," 2001, http://www.users.waitrose.com/~daviddewhurst/Projects.html.

[5] P. Picton and M. Capp, The Optophone: Some question and answer, School of Technology and Design. Nene - University College Northampton, 2000, http://oldweb.northampton.ac.uk/aps/eng/research/optophone/optophone.html.

[6] G. Sainarayanan, R. Nagarajan, and S. Yaacob, "Vision substitution for visually handicapped by stereo acoustic of image using neural network," in Proc. International Conference on Engineering Application of Neural Networks (EANN '00), pp. 281–287, Cagliari, Italy, July 2000.

[7] J. R. Parker, *Algorithms for Image Processing and Computer Vision*, John Wiley & Sons, New York, NY, USA, 1997.

[8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, Upper Saddle River, NJ, USA, 2nd edition, 2002.

[9] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 6, pp. 679–698, 1986.

[10] P. Garcia, F. Pla, and I. Garcia, "Detecting edges in colour images using dichromatic differences," in *Proc. 7th International Conference on Image Processing and Its Applications*, vol. 1, pp. 363–367, Manchester, UK, July 1999, (Conf. Publ. no. 465).

[11] G. Bellaire, K. Talmi, E. Oezguer, and A. Koschan, "Object recognition: obtaining 2-D reconstructions from color edges," in *Proc. IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI '98)*, pp. 192–197, Tucson, Ariz, USA, April 1998.

[12] E. Saber, A. M. Tekalp, and G. Bozdagi, "Fusion of color and edge information for improved segmentation and edge linking," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '96)*, vol. 4, pp. 2176–2179, Atlanta, Ga, USA, May 1996.

[13] S. Wesolkowski, M. E. Jernigan, and R. D. Dony, "Comparison of color image edge detectors in multiple color spaces," in *Proc. IEEE International Conference on Image Processing (ICIP '00)*, vol. 2, pp. 796–799, Vancouver, BC, Canada, September 2000.

[14] The MathWorks Inc., Image Processing Toolbox, User's Guide Version 3, 2001.

[15] J. C. Russ, *The Image Processing Handbook*, CRC Press, Boca Raton, Fla, USA, 2nd edition, 1995.

[16] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*, McGraw-Hill, New York, NY, USA, 1995.

[17] C. Capelle, C. Trullemans, P. Arno, and C. Veraart, "A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 10, pp. 1279–1293, 1998.

**R. Nagarajan** received his B.E. degree (honors) in electrical and electronics engineering from Madras University in 1961, the M. Tech. degree in control engineering from IIT Kanpur in 1969, and the Ph.D. degree in adaptive control from Madras University in 1977. Currently he is a Professor in the School of Engineering and Information Technology, Universiti Malaysia Sabah, Malaysia. His current research interests include robot control and intelligent process control. He is a Life Fellow of Institution Engineers, India, a Senior Member in IEEE, USA, and a Member of UICEE, Australia.

**G. Sainarayanan** is currently a Lecturer in electrical and electronics engineering in Universiti Malaysia Sabah, Malaysia. He received his B.E., M.E., and Ph.D. degrees, respectively, from Annamali University, India, Bharathiar University, India, and Universiti Malaysia Sabah, Malaysia, in 1998, 2000, and 2002. His research interests are in the areas of vision rehabilitation, medical imaging, and intelligent control.

**Sazali Yaacob** received his B.E. degree in electrical engineering from Universiti Malaya in 1986, the M.S. degree in system engineering from the University of Surrey in 1987, and Ph.D. degree in control engineering from the University of Sheffield in 1995. Currently he is a Professor in the School of Mechatronic Engineering, Northern Malaysia University College of Engineering, Malaysia. His research interests include acoustics, modeling, and control. He was granted a Charted Engineer status by the Engineering Council, United Kingdom, in 2004.

**Rosalyn R. Porle** received her B.E. degree (honors) in electronics and telecommunication engineering from Universiti Malaysia Sarawak (UNIMAS), Malaysia, in 2001. Currently, she is a Master student in the electrical and electronics program, School of Engineering and Information Technology, Universiti Malaysia Sabah (UMS), Malaysia. Her research interests are in the fields of colour image processing, fuzzy logic, and neural network.