Rehabilitation of the Blind using Audio to Visual Conversion Tool

Ashwani Kumar Aggarwal

Electrical and Instrumentation Engineering Department, SLIET, Longowal, Sangrur, Punjab, INDIA

ashwani_75@yahoo.com

ABSTRACT

In human beings, eyes play a vital role. A very less research has been done for rehabilitation of blindness for the blind people. This paper discusses the work that helps blind people for recognizing the basic shapes of the objects like circle, square, triangle, horizontal lines, vertical lines, diagonal lines and the wave forms like sinusoidal, square, triangular etc. This is largely achieved by using digital camera, which is used to capture the visual information present in front of the blind person and a software program, which achieves the image processing operations, and finally the processed image is converted into sound. After the sound generation process, the generated sound is fed to the blind person through headphones for visualizing the imaginary image of the object. For visualizing the imaginary image of the object, it needs to train the blind person. Various training process methods had been applied for recognizing the object.

Keywords: Image processing, Pixel, Pitch, Loudness, sound generation, Edge detection, brightness

1. INTRODUCTION

In human beings Eyes will play vital role. Generally the vision impairment caused due to several reasons such as blindness by birth, various damages to the brain etc[1]. In order to help blind people various instruments had been developed for rehabilitation of blindness [2]. But still the developed instruments are not satisfactory for the rehabilitation of blindness [3]. The present work aims at the rehabilitation of the blind person to recognize the basic shapes of the objects, because in the universe every object has some shape that is the combination of the basic shapes like Circle, Square, Triangle, Horizontal lines, vertical lines, Diagonal lines etc. This is achieved by, the visual information present in front of the blind person is captured by using a digital camera, later the acquired image is undergone various image processing techniques and the processed image is converted into sound. By listening to the sound patterns the blind

person has to visualize the imaginary image of the original object. The block diagram of the system and sound generation process is explained in the following sections of this paper. After the sound generation process, it needs to train the blind person for recognizing the objects. After training the blind person, more than 80% of the results are satisfactory. Before going to the inner details of the work, the fundamentals of the digital image and digital audio have been discussed.

1.1   Fundamentals of a Digital Image

A Digital Image is defined as rectangular array of dots called Pixels [4]. Each digital image is comprised of many pixels and each pixel has its own color. High quality Image in general requires more pixels than low quality images. Generally digital images can be produced as Black & White, Grayscale and color. The color information in the pixel is represented in terms of the Bit depth of an image. Bit depth is defined as the number of bits used to define each pixel[5].

1.2   Introduction to Digital Audio

Sound is what we experience when the ear reacts to a certain range of vibrations, such as a guitar string. As the string moves in one direction, it pushes on nearby air molecules, causing them to move closer together. This creates a small region of high pressure on one side of the string and low pressure on the opposite side. As the string moves in opposite direction, the areas of high and low reverse [6].
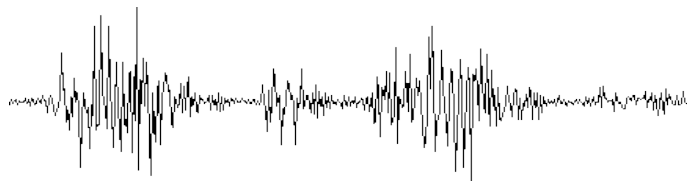

Figure 1.2 A sample Analog Sound Signal

The Sono Visualization aims the blind person to recognize the basic shapes of objects like Horizontal Lines, Vertical lines, diagonal lines, circle, square, triangle etc by listening sound, because in the universe every object has some shape that is the combination of basic shapes. By listening sound patterns the blind person has to visualize the imaginary image of the original object present in the image [7]. For visualizing the imaginary shapes of the objects it needs to train the blind person.

2. IMAGE TO SOUND CONVERSION

Different sinusoidal oscillators form an orthogonal basis, assuming they have frequencies that are integer multiples of some basic frequency [8]. A pixel at a more elevated position i corresponds to an oscillator of higher frequency f_i. The larger the brightness of a pixel represented by gray tone , the larger the amplitude ("loudness") of its associated oscillator [9].
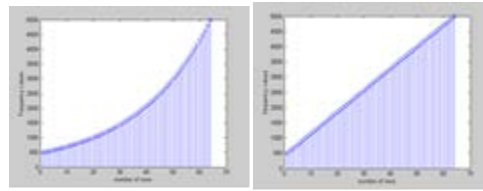
The M oscillator signals for the single column are superimposed, and the corresponding sound pattern s(t) is presented to the ear during T/N seconds. Then the next, (j+1)th, column, is converted into sound: this procedure continues until the rightmost, Nth, column has been converted into sound, T seconds after the start of the conversion.   For a given column j, every pixel in this column is used to excite an associated sinusoidal oscillator in the audible frequency range. For 64 x 64 pixel matrix, M = 64 behavior of sinusoidal oscillators are implemented digitally, whose frequency values of sinusoidal waves is calculated by using exponential or Equidistant frequency distributions, amplitude of the sine waves is calculated by using color (brightness) of the pixels and the phase values are generated randomly in the range of 0 to 1. The phases   are just arbitrary constants during the image to sound conversion process [10]. The ear is insensitive to phase because it contains little useful information. But we cannot say that the ear is completely deaf to phase because a phase change can rearrange the time sequence of an audio signal [11]. Once in a column after processing of the entire pixel information then mix all the signals produced in a column. Then increment the column and repeat the process until the last column. All the sound samples, which are generated, are written to the wave file by adding the header [12].

3. RESULTS

In the Sono-Visualization system, all the algorithms were developed in C language by using VC++ Compiler 6.0. The input color JPEG images were generated either with Microsoft paint or by using Logitech web camera, which was then converted into BMP grayscale image [13]. Some of the input images and the corresponding output images are shown below. Once after the sound generation, the sound was fed to the blind person through headphones.

Table 1: Results of Horizontal, Vertical and Diagonal Lines

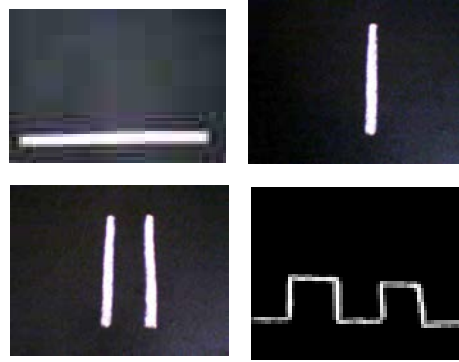| Time of the sound file Played in sec	1 horizontal line	2 horizontal lines	3 horizontal lines
2.10	Y	N	N
5.25	Y	N	N
10.50	Y	Y	N

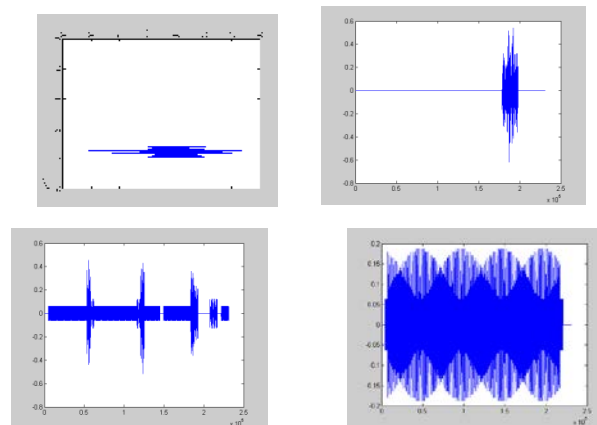**Figure 1: The input image containing a Square wave**

Figure 2: Output wav sound file graph.

For recognizing the object it needs to train the blind person [14]. Various procedures have been applied for training the blind person. The present work had chosen the blind persons with the age in the range of 15-20 years. First the blind person was trained with horizontal, vertical, diagonal lines and then to square, circle, triangle etc and then to other complicated structures. The results are shown in the tabulated form. If the blind person recognized the object then it is represented as 'Y' stands for 'Yes' otherwise 'N' for 'No'

The blind person is able to recognize the vertical lines in short period of time. The blind person is able to recognize the vertical lines in minimum period of time [15].

The blind person was able to recognize the single horizontal line in an image, but in case of 2 to 3 horizontal lines, the person told some multiple lines were there so he was not able to trace the exact number of ones. In minimum time period the blind person was not able to trace the both the up & down diagonal lines. But with increase of time the person had recognized both the shapes.

4. COMMENTS AND DISCUSSION

Here first the blind person was trained with the various basic shapes and the random shapes of sound files are played for recognition. In minimum time period the blind person was

in confusion to recognize the object, but with increase of time the person had recognized successfully. With the available feedback, the blind persons were then trained with the combination of the above mentioned shapes like combination of horizontal and diagonal, diagonal and vertical etc. Here also the results are satisfactory, they recognized the combination of the shapes also.

In next phase the blind persons were trained with the waveforms like sine waves, triangle waves, square pulses, saw tooth waves etc. This had taken more time to recognize the wav forms. But the results still are satisfactory.

The blind persons visualized the vertical lines easily and they were able to tell how many vertical lines were there in the image. But in case of horizontal lines, they were able to recognize the horizontal line but not able to recognize how many lines were present. Even the sound file playing time was increased to 10 seconds; they were hardly able to recognize the number of lines. But more than 80% of the results are satisfactory. Even the combination of the shapes was also recognized. For more effective recognition if we generate the stereo sound, it would help the blind person to recognize the direction of the object i.e. whether it is there on right or left. For recognizing the complicated shapes it needs a lot of training and the blind person has to use this tool daily.

5. CONCLUSION

The Sono Visualization Blind Aid Tool is developed to help blind people in recognizing the basic shapes of the objects like circle, square, rectangle, horizontal, vertical and diagonal lines etc. The blind persons trained with these basic shapes, with less effort that they easily recognized the basic shapes. With continuous training they can be able to recognize the multiple objects and combination of the objects. The main important aspect in this is the sound generation process. The sound is generated by simple mapping of the image properties into sound properties. Sound is generated in the form of simple sinusoidal wave forms. The mono type sound is generated by adding the wav header file to the sound samples for playing the sound file. Wav file format is used because it is the windows operating system most favorable format. Edge detection of the image and conversion of the JPEG to BMP image conversion are provided as an optional for user convenience [16]. Edge detection helps the blind person to recognize the solid shapes of the objects. JPEG to BMP conversion procedure is used for user compatibility to work with both JPEG and BMP image files[17].

6. FUTURE SCOPE OF THE WORK

6.1 Generation of Stereo sound

The present work generates the Mono type sound form; in advanced version to recognize the objects more effectively stereo sound is generated. Stereo sound helps the blind person to recognize the objects easily [18] whether it is there on right or left. If the object is there on right, the generated sound will first reach the right ear and vice versa. The stereo sound is generated in wav file format with 16-bit resolution [19]. Scanning is simultaneously carried out on the first column and last column of the image. After this scanning procedure, the generated sound is outputted to the headphones. The procedure is implemented until the centerline of the image is met. Sound will be produced after every simultaneous scanning of the two-mirrored columns in each image[20].

6.2 Developing of prototype and Mobile based system

In order to work in the real time environment a prototype system is developed, which consists of a head mounted digital camera, a suitable DSP system to achieve image processing and sound generation process are assembled with the headphones**.** The system should be portable, should be run on battery operated and should have low power consumption capability[21]. The system can be developed on a Mobile based device as the Digital cameras with built in camera are available easily in the market at low cost. The visual information present in front of the blind person is captured by using built in camera of a mobile phone and a suitable software program generates the sound. The generated sound is then fed to the blind person through head phones to visualize the imaginary image of the original object. The software program is made to run by using either Symbian or Nucleus operating system [22].

6.3 Encoding of the sound in the form of MIDI and MP3

In general it is difficult to recognize the normal sounds for reorganization of the objects. To overcome this problem the normal sounds are encoded in the form of MIDI (Musical Instrument for device interfacing). The human beings can easily grasp the music compared to the normal sounds. For different objects different sounds of instruments are to be encoded, for example if we encode the piano sound for an object, if the object is located at high then the intensity of the piano music is more and vice versa. Encoding the sound into music helps the blind person to recognize the object more easily.The present work generates the sound in the form of wav file format. But the disadvantage with the wav files is the size of the file is very large and it is in the uncompressed format. It will occupy more memory space, this is a major problem when the system is developing on mobile based, because in the mobile based devices have the limited memory. To overcome this problem the sound is generated in the form of MP3, which is compressed format and the size of the file is very less. The MP3 sound is generated by using Wav to MP3 encoding procedure [23].

6.4 Further future scope of the work

The further future scope of this work aims the blind person to recognize the color of the object, whether the object is red or green or blue, size of the object whether it is large or small. For reading the text like alphabets and numbers is also achieved with the help of Optical character recognition (OCR) machine, to recognize the characters. After that by using a speech synthesizer it will say the name of the alphabet or number like A, C, D or 1, 2, 3. The output of the speech synthesizer is fed to the headphones of the blind person for reading the text.

REFERENCES

[1].   Gonzalez R.C. and Woods Richrd. E "Digital image processing" second edition, published by Pearson education. Pte.led., pp 54-62, March-2003.

[2].   B.Thylefors, A.-D. Negrel, R. Paramjasegaram and K.Y. Dadzie, " Global Data on blindness", *Bulletin of the world health organization* 73(1), pp.115-121,1995.

[3].   Tohru Ifukube, Tadayuki Sasaki and Chen Peng, "A Blind Mobility Aid Modeled after echo-location of Bats" *IEEE Trans. Biomed. Eng.,* Vol.38, pp.461-465, May 1991.

[4].   S. Shoval, J. Borestein and Y. Koren, " The NAVBELT – a computerized Travel Aid for blind Based on Mobile Robotics Technology", *IEEE Transactions on Biomedical Engneering*Vol.45, No.11,pp. 1376-1386, November 1998.

[5].   F.Wong, R.Nagarajan, S.Yaacob, A.Chekima and N-E. Belkhamza, " Electronic Travel Aids for Visually Impaired – a Guided Tour", proceedings of conference in Engineering in Saraw-ark In the Last century and for the Next Decade, Kuching, Sarawak, pp. 111-129, May 2000.

[6].   L.Kay, " A sonar aid to enchance spatial perception of the blind: Engineering, design and evaluation," Radio, Electron, Engineer, vol.44, pp.605-627, 1974.

[7].   Brabyn, J., Crandall, W., Gerrey, W., "Talking signs: a remote signage, solution for the blind,visually impaired and reading disabled" Engineering in Medicine and Biology Society, Proc-eedings of the 15th Annual International Conference of the *IEEE* Oct 28- 31, 1993 Page(S): 1309 – 1310.

[8].   C. Capelle, C.Trullemans, P.Arno and C.veraart, "A Real Time Experimental prototype for Enhancement of vision rehabilitation using Auditory Substitution", *IEEE Transactions onBiomedical Engineering,* vol.45,No.10, pp 1279-1293, October 1998.

[9].   Meijer, P.B.L., "An Experimental System for Auditory Image Representations"*, IEEE Trans. on Biomedical Engineering,* Vol.39, No.2, pp.112-121, February 1992.

[10].   Wm. H. Dobelle, "Artificial Vision for the Blind by Connecting a Television Camera to the Visual Cortex" *ASAIO Journal 2000; 46:3-9.*

[11].   Cronly-Dillon, J.,Persaud, K. C., Blore, R., (2000) "Blind subjects construct conscious men tal images of visual scenes encoded in musical form." *Proceedings Of The Royal Society London Series B-Biological Sciences*, *267,* 2231-2238.

[12].   Matta, S., Kumar D. K., Xinghuo Yu, Burry M, "Discriminative Analysis for Image to sou- d Mapping" Intelligent sensing and information processing, 2004. Proceedings of internat- ional conference on 2004 page(s);119-122.

[13].   J. Dowling, A. Maeder and W. Boles, ``Intelligent Image Processing Constraints for Blind Mobility Facilitated through Artificial Vision,'' *Eighth Australian and New Zealand Inteli gent Information Systems Conference* (ANZIIS 2003), Sydney, Australia, pp. 109-114, December 10-12, 2003.

[14].   R. Nagarajan, S. Yaacob and G. Sainarayanan, ``Role Of Object Identification In sonofication System For Visually Impaired,'' *IEEE Tencon 2003* ("IEEE Region 10 Conference on Convergent Technologies for the Asia-Pacific"), October 15-17, 2003, Bangalore, India.

[15].   Gonzalez R. C. and Woods R. E "Digital image processing" second edition, published by Pearson education.Pte.led., pp 38-42, March-2003.

[16].   W.F.Ganong, *Review of Medical Physiology*, 15th ed., Appleton & Lange, Eds. London, U.K.: Prentice-Hall 1991.

[17].   E. Zwicker, "Scaling", in handbook of sensory Physiology, vol, V/2 Auditory system, phy-siology (CNS), Behavioral studies, psychoacoustics, W. D. Keidal and W. D. Neff, Eds. Berlin,Germany: Springer-verlag, 1975.

[18].   Steven W. Smith "The scientist and Engineer's Guide to Digital signal processing" second edition published by California Technical Publishing, pp.351-370., chapter 22 –Audio Proc.

[19].   G. K. Wallace, "The JPEG still picture compression standard", *Communications of the   ACM*, vol.34, no.4, pp.30-44, Apr.1991.

[20].   Wong, F.; Nagarajan, R.; Yaacob, S.; Chekima, A.; Belkhamza, N.-E.;"A stereo auditory display for visually impaired" TENCON 2000. Proceedings Volume 2, 24 -27   Sept.,2000.,Page(s):377-382,vol.2