International Journal of Advanced Technology in Engineering and Science
Vol. No.4, Issue No. 09, September 2016
www.ijates.com
ijates
ISSN 2348 - 7550

PRE-PROCESSING ON STEREO IMAGES FOR COMPUTER VISION BASED TRAVEL AID FOR BLIND

Shrugal Varde[1], Prof. M.S. Panse[2]

[1] *PhD. Scholar, Electronics Department, VJTI, Matunga,Mumbai (India)*

[2] *Professor and Dean Faculty, Electronics Department, VJTI, Matunga, Mumbai (India)*

ABSTRACT

Electronic travel aid, which convert visual environmental cues into another sensory modality, have proved to help blind people with a great degree of physiological comfort and independence. Designing an obstacle detector for blind is a challenging task. Most of these electronic travel aids available today are based on ultrasound, infra-red and sonar technology and a very few travel aids are based on stereo image processing. Stereo imaging as a substitute for human vision is a powerful tools with great potential to enable a range of assistive technologies for the growing population of blind and visually impaired people. But images acquired by stereo cameras or other imaging systems cannot be used directly as they may be corrupted by random variations in intensity, variations in illumination, or poor contrast. To increase the accuracy of depth estimation of obstacles in front of a visually impaired person, pre-processing of the stereo images is required. This paper describes the pre-processing tasks like histogram equalization, image filtering and edge detection to be performed on stereo images and gives a comparative study of time required for each pre-processing task on images with different resolutions.

Keywords: Electronic Travel Aid, Stereo Imaging, Edge Detection, Histogram, Bilateral Filter, Visually Impaired.

I. INTRODUCTION

Mobility means moving safely, gracefully and comfortably [2]. It relies in large part in perceiving properties of immediate surroundings and avoiding obstacles, negotiating steps, drop-off etc. Information of the environment enables humans to know various sources that are in many different directions. Proper source localization is very important for safety, survival and navigation.

Vision loss affects almost every activity of daily living. According to global data on visually impaired report submitted by world health organization, in world there are about 289 million people blind [1]. This accounts to almost 4.7% of total world population. Out of 289 million people 246 million people have low vision, that is 86% of total blind population and 39 million people are completely blind. As per this report in India, out of 1181.4 million population there 8.075 million people completely blind, 54.544 million people with low vision and 62.619 million people visually impaired. These figure clearly indicates that there should be a mobility aid that can help visually impaired people.

There are many mobility devices and navigation aids developed for visually impaired population. Most commonly used are white cane and guide dog. Commonly used electronic travel aids includes navblet, guidecane, and many more. All the available electronic travel aids are based on ultrasound or sonar technology. These technologies have a few drawbacks, they cannot cover surface to head height obstacles, more number of sensors are required to scan a larger environment, and training time required is more.

Computer vision and mobile computing as a substitute for human vision are powerful tools with great potential to enable a range of assistive technologies for the growing population of blind and visually impaired people. Applications based on computer vision enhance these persons' mobility and orientation as well as object recognition and access to printed material. Stereo depth map technique is method used to detect depth of objects captured by two cameras.

Block diagram of prototype of electronic travel aid based on stereo depth map algorithm is as shown 1.

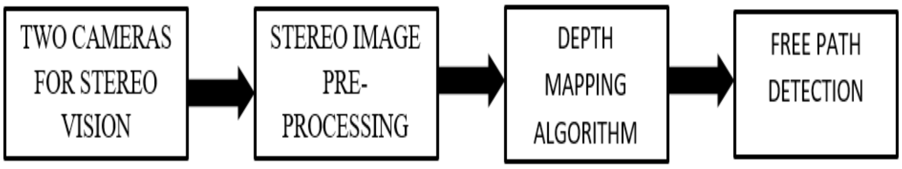

Figure 1: Block Diagram of Electronic Travel Aid Based on Stereo Depth Map Technique

There are two cameras that capture images at a pre-defined rate [3]. Stereo depth map algorithm is used to calculate disparity of each pixel, and from disparity the relative distance of obstacle from the user. The captured image cannot be directly used to calculate disparity as they might be corrupted by random variations in intensity, variations in illumination, noise, or poor contrast. To reduce the effect of noise and difference in contrast ratio of two images and to increase the accuracy of stereo depth map algorithm, some pre-processing on the captured images is required. Block diagram of pre-processing block is as shown in figure 2.

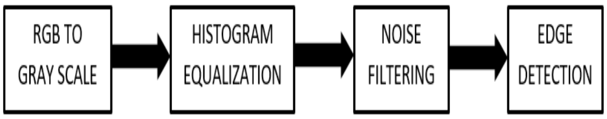

Figure 2: Pre-processing block

Section II to section V explains the preprocessing tasks performed on captured images. Section VI explains the results and future scope of the design.

II. GRAY SCALE CONVERSION

Camera modules used in the proposed design provides 8 bit RGB data. Hence the acquired image contains three matrices, one for red, one for green and one for blue intensity. Each pixel of these matrices is of 8 bits. If depth map algorithm was directly applied to colour images, then algorithm had to process three separate matrix thus increasing the processing time. The main task of depth map algorithm is to extract depth information and there is no colour segmentation of images in stereo algorithm, hence the color information is not important. Hence we convert the color image to intensity image or gray scale imagee. We use the luminosity method to convert RGB to gray scale image [4]. This method is much more accurate than simple average method. The formula to get intensity is given as

$$L = 0.21 * R + 0.72 * G + 0.07 * B \ldots (1)$$

Where R is the red matrix, G is the green matrix and B is the blue matrix and L is the gray scale image.

Figure 4: Captured RGB image from left

Figure 3: Gray scale image

III. HISTOGRAM EQUALIZATION

Histogram of a digital image with intensity levels in the range [0, L-1] is a discrete function $h(r_k) = n_k$ where r_k is the kth intensity value and n_k is the number of pixels in the image with intensity r_k. Sometimes same scene captured from two locations separated by a small distance can lead to the particular point in the three dimensional world imaged by two different intensities. In depth map algorithm we find the correlation between the pixels of stereo images. Variation in intensities for the same point can lead to wrong pairing of pixels and hence it is very important to perform histogram matching before using the depth map algorithm [5].

Discrete histogram equalization transformation is given as

$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) \ldots (2)$$

Since

$$p_r(r_j) = \frac{n_k}{MN} \ldots (3)$$

Where M and N are number of rows and columns of grey scale image.

Hence

$$s_k = T(r_k) = \frac{(L-1)}{MN}\sum_{j=0}^{k} n_k \ldots (4)$$

Given a specific value of s_k the discrete formulation includes computing transformation function as

$$G(z_q) = (L-1)\sum_{i=0}^{q} p_z(z_i) \ldots (5)$$

for values of q so that

$$G(z_q) = s_k \ldots (6)$$

$p_z(z_i)$ is the ith value of specified histogram. Hence we get z_q as

$$(z_q) = G^{-1}(s_k) \ldots (7)$$

The implementation of the above equations (4) to (7) is explained in figure 5

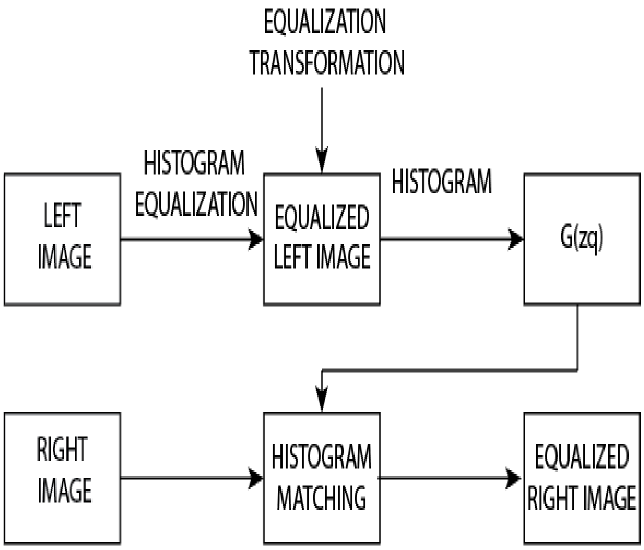

Figure 5: Block diagram of histogram equalization

Results of histogram equalization process are as shown in figure 6 and figure 7

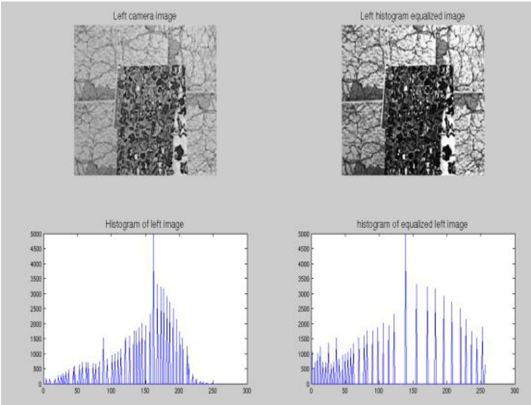

Figure 7: Left image histogram equalized

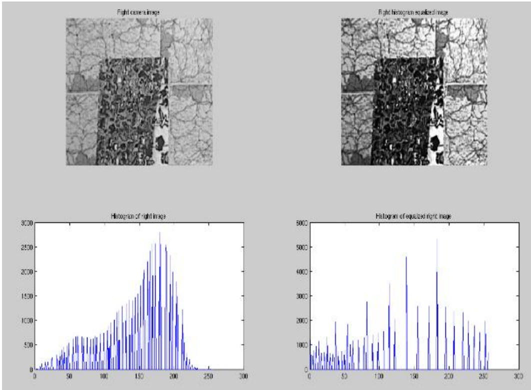

Figure 6: Right image histogram equalized

IV. IMAGE NOISE FILTERING

Images are often corrupted by random variations in intensity values, called noise. Some common types of noise are salt and pepper noise, impulse noise, and Gaussian noise. Salt and pepper noise contains random occurrences of both black and white intensity values. Impulse noise contains only random occurrences of white intensity values. Unlike these, Gaussian noise contains variations in intensity that are drawn from a Gaussian or normal distribution and is a very good model for many kinds of sensor noise. Linear smoothing filters are good filters for removing Gaussian noise and, in most cases, the other types of noise as well. A linear filter is implemented using the weighted sum of the pixels in successive windows. Typically, the same pattern of weights is used in each window, which means that the linear filter is spatially invariant and can be implemented using a convolution mask. Most commonly used linear filter to remove noise from images in the Gaussian filter [6]. Gaussian filter is also called as blur filter. Gaussian blur filtered image is given by

$$GB[I_p] = \sum_{(q \in s)} G_\sigma(\|p - q\|) I_q \ldots (8)$$

Where

I_p - Window of image.

p - centre pixel of the window

q – Neighbouring pixels of the window

and

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right) \ldots (9)$$

is a two dimensional Gaussian kernel.

Effect of σ on the image is as shown in figure 8

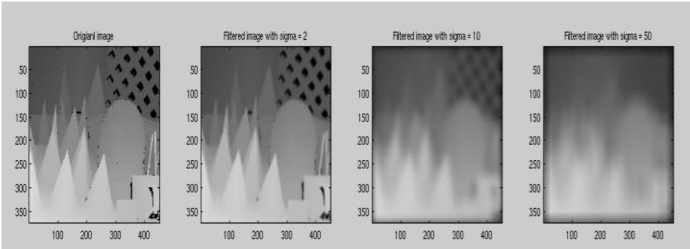

Figure 8: Effect of sigma on image

From figure 8 we can say that as the value of σ increases, the information about the edges is lost. To overcome this problem we use bilateral filter. Bilateral filter is non-linear filter which is the modification of Gaussian filter [7]. Similarly to the Gaussian filter, the bilateral filter is also defined as a weighted average of pixels. The difference is that the bilateral filter takes into

Account the variation of intensities to preserve edges. The rationale of bilateral filtering is that two pixels are close to each other not only if they occupy nearby spatial locations but also if they have some similarity in the photometric range.

The bilateral filter, denoted by BF[.] is given as

$$GB[I_p] = \frac{1}{Wp} \sum_{(q \in s)} G_{\sigma s}(\|p-q\|) G_{\sigma}(I_p - I_q) I_q \ldots (9)$$

Where Wp is given as

$$Wp = \sum_{(q \in s)}' G_{\sigma s}(\|p-q\|) G_{\sigma r}(I_p - I_q) \ldots (10)$$

Parameters σ_s and σ_r will measure the amount of filtering in image I. $G\sigma_s$ is a spatial Gaussian that decreases the influence of distant pixels, $G\sigma_r$ a range Gaussian that decreases the influence of pixels q with an intensity value different from Ip.

Figure 9 shows that bilateral filter was able to reduce the noise and retain the edges information.

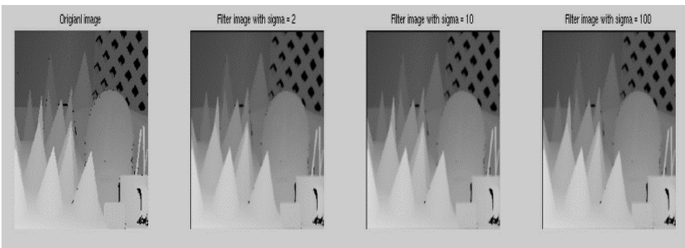

Figure 9: Effect of Bilateral filter

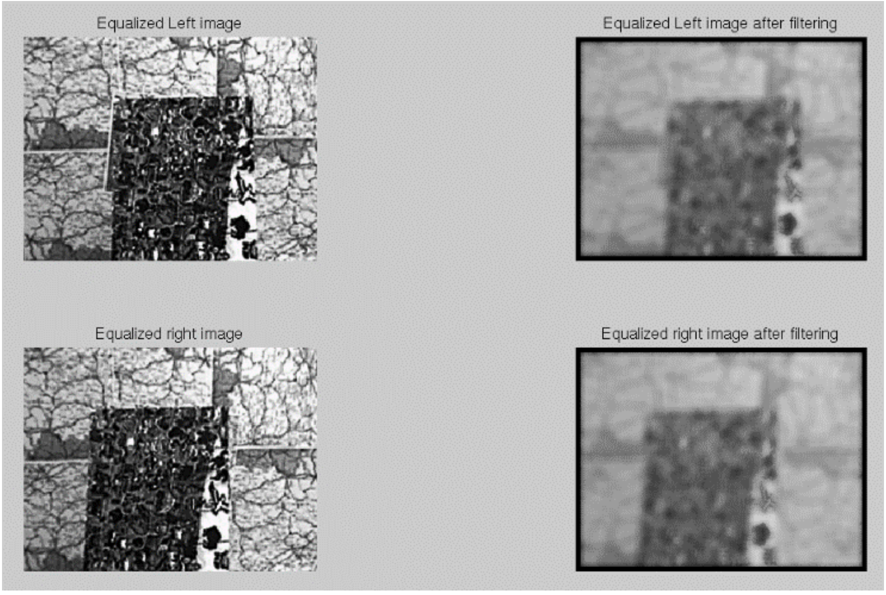

Figure 10 : Bilateral filter applied to left and right camera image

V. EDGE DETECTION

Edge detection refers to the process of identifying and locating sharp discontinuities in an image. The discontinuities are abrupt changes in pixel intensity which characterize boundaries of objects in a scene [8]. Classical methods of edge detection involve convolving the image with an operator (a 2-D filter), which is constructed to be sensitive to large gradients in the image while returning values of zero in uniform regions. There are an extremely large number of edge detection operators available, each designed to be sensitive to certain types of edges. The geometry of the operator determines a characteristic direction in which it is most sensitive to edges. Operators can be optimized to look for horizontal, vertical, or diagonal edges. There are five main edge detection algorithms. They are classical filters (Sobel, prewitt...), zero crossing algorithms, laplacian of gaussian (LoG), canny edge detection.

Classical filter like sobel, prewitt are simple to implement but are very sensitive to noise. These filters were designed to detect edges along the horizontal and vertical direction. Zero crossing edge detector, like classical filter are sensitive to noise. Laplacian of Gaussian (LoG) is great method to find the exact location of edges, but his filter cannot find the orientation of edges because of the use of laplacian filter. Canny edges detector uses probability method to detect the edges. This filter can detect edges in any orientation and has better detection of edges in noisy conditions. Only disadvantage of this method is because of its complexity implementation time is more. Taking into consideration advantages and disadvantages of the various filters and also the implementation results of filters on same image, canny edge detector filter was used in the design. The implementation of various filters on the same image is as shown in figure 11.

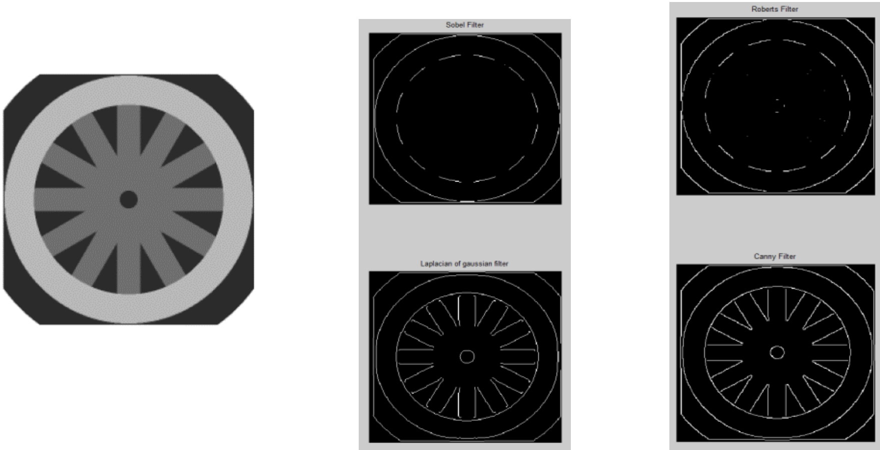

Figure 11: Edge detector comparison

Output of canny edge detector for stereo image pair used is as shown

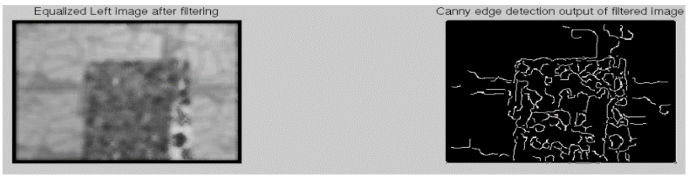

Figure 12: Left camera image edge detection

VI. RESULTS AND CONCLUSION

The implementation of pre-processing block was done in Matlab 2014b simulation software. Hardware description of the system is i5 Intel fourth generation dual core processor with 8 GB ram. Linux operating system was used. For bilateral filter, the size of the window used was 15*15, value of σ_s was selected as 0.5 and σ_r was selected as 0.04. Time required for various stages in pre-processing tasks for different resolution of images is as shown in table below.

	Time taken by each pre-processing stages(in seconds)			
Resolution of images	RGB to grey scale	Histogram equalize	Image filter	Edge Detector
1110*1240	0.003	0.072	12.05	1.14
999*1116	0.0036	0.0635	11.41	0.901
888*992	0.003	0.0535	10.104	0.82
777*868	0.003	0.057	9.093	0.6805
666*774	0.003	0.06	7.852	0.646

Table 1: Comparison table of time consumed by pre-processing stages

From the above table it is clear that for any resolution images, the time consumed by the RGB to gray scale, histogram equalization step, and edge detector stage is less than 1 second. The only time consuming step in the complete pre-processing task is bilateral filter stage. As the resolution of images increases, the time consumed by this stage also increases. To implement real time obstacle detection system for visually impaired people, modification in the bilateral filter algorithm to reduce the total time consumption is required.

REFERENCES

[1] Global data on visually impaired, World health organization, 2010.

[2] Dimitrios Dakopoulos, Nikolaos Bourbakis," Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol.40, no.1, pp.25-35, Jan.2010.

[3] Shrugal Varde, Prof. M.S.Panse, "Computer vision based travel aid for blind", International journal of latest technology in engineering, management & applied science, vol. 4,no. 8,pp. 58-63,Aug 2015.

[4] C. Saravanan, "Color Image to Grayscale Image Conversion," Second International Conference on Computer
Engineering and Applications (ICCEA), Bali Island, 2010, pp. 196-199.

[5] R. C. Gonzalez, Digital image processing. Perarson Education, 2009.

[6] M. Zhang and B. K. Gunturk, "Multiresolution Bilateral Filtering for Image Denoising," IEEE Transactions on Image Processing, vol. 17, no. 12, pp. 2324-2333, Dec. 2008

[7] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," Sixth International Conference on Computer Vision, 1998. Bombay, 1998, pp. 839-846.

[8] M.C. Shin, D. Goldgof, and K.W. Bowyer. "Comparison of Edge Detector Performance through Use in an Object Recognition Task". Computer Vision and Image Understanding, vol. 84, no. 1, pp.160-178, Oct. 2001