Indian J. Innovations Dev., Vol. 1, No. 8 (August 2012) ISSN 2277 - 5390

Appliance control using human detection and disparity estimation

***K. Narasimhan**

School of Electrical & Electronics Engineering
Department of Electronics & Communication Engineering
SASTRA University, Thanjavur, Tamilnadu, India.

****Corresponding author**: Dr. K. Narasimhan, Assistant Professor School of Electrical & Electronics Engineering, Department of Electronics & Communication Engineering SASTRA University, Thanjavur, Tamilnadu, India.*

. **Abstract**

We here propose a novel method for detecting and locating the human presence for controlling the power appliances in the subjected area. At the heart of our Endeavour, we use an over-head camera to capture the scene inside the concerned arena. From the scene we sample frames and process them to detect silhouette of objects. We then use a codebook for instance as a reference to filter out human outline(s) to detect human. Their location in the room is traced using disparity calculation and subsequent depth extraction. Once the location has been estimated, the corresponding appliances can be operated accordingly for service. A controller interfacing can be provided to achieve the above. This is done using the serial communication technique with RS 232. Presence of max 232 IC guides the above. The serial communication is performed by specifying the port and baud rate of use and burning the needed code to perform the task.

Keywords: Silhouette, Code book, Disparity calculation, Depth Estimation, Microcontroller interface, Serial Communication, RS232.

Introduction

It all started with the carelessness of the people who tend to sway away without caring to turn off the electrical devices that were used up. Adding, the most haunting problem of today is the shortage of power, also leading to increase in cost per unit. The objective of our work involves human detection from video which involves silhouette estimation, code book generation to distinguish human and clutter, disparity estimation, depth extraction to locate 3-dimensional position of human, serially communicating the 3-dimensional location to microcontroller, which in tum operate relay for the control of switches in the concerned area. For human detection various algorithms are used with respect to surveillance application, human motion detection, face recognition etc. Seyyed Meysam Hosseini *et al.*, 2010, proposed a novel method of new human gait recognition system in which feature extraction is done based on radon transform of binary silhouette. Multilayer feed forward neural network is used to make final decision. Spatio temporal gait curve method is used by Brian & Arun 2010, in surveillance application for the task of object detection and human identification during the night time. Performance of these algorithms is verified with CASIA Night gait database. Background modeling based human motion detection and analysis for video surveillance analysis to analyze people activities such as walking, running is proposed by Murat EK, 2005. Processing interface difference image being capable of segmenting human silhouette from dynamic scene is used by Piotr & Pawel, 2006). Template matching based supervised learning method is used to distinguish human, human group and human action into predefined classes like walking, boxing etc is proposed by (Yigithan *et al.*, 2006). Cyclic gait analysis used to extract key frames which are compared with training frames using normalized correlation and subject classification done by the nearest neighbor classifier is used by (Robert T.Colin *et al.*, 2002). In literature various algorithm like Windows method, Global Optimization method, Belief propagation, Max field, block-matching, pel recursive, optical flow, Bayesian approach, Graph cut etc are used for disparity estimation. Novel two

Indian J. Innovations Dev., Vol. 1, No. 8 (August 2012) ISSN 2277 - 5390

stage disparity estimation algorithm namely Iterative Block and Octagonal Matching (IBOM) is proposed by Shao-Yi Chien *et al.,* 2003. IBOM algorithm gives good result and eighteen times faster than Octagonal matching algorithm. By minimizing a differentiable cost function, high precision real valued disparity estimation is done with Hidden Markov Random Measure field Model (HMMF), by (Edgar Arce *et al.,* 2007).Curve fitting tool is used by (Anwar hasni *et al.,* 2009) for the navigation of stereo vision autonomous vehicle. Dense disparity map is produced by minimization of energy function with the help of simulated annealing, multi-resolution and interpolation. Temporal modeling using scene flow is proposed by (Fang Liu *et al*., 2008) which improves temporal consistency in the disparity map. Image shifting and overlapping area to estimate the disparity value is proposed by Yen San Yong & Hock Woon Hon, 2008. Geometrical relation between camera and model point pair is used for depth estimation approach by Zhiguang Zhong *et al.,* 2005. Hardware implementation of maximum likelihood stereo correspondence algorithm on FPGA is proposed by Siraj Sabihuddin & W. James MacLean, 2007. Depth estimation obtained by using stereo matching is further corrected by sparsity of the original depth image is proposed by Sakuragi, et al. 2010.

The work has been divided into three phases for easier implementation as follows,

1. Human detection
2. Disparity estimation
3. Microcontroller interface

Now the result obtained from these three phases are combined together to achieve the objective. Here, the results obtained from each of these phases are analyzed individually for accomplishing our motto.

Materials & Methods

Materials Required

- The camera used is an 8 MP I-ball web camera. It is mounted on a wooden case as such to ensure stability, MATLAB 7.6.0 (R 2008a), Phillips microcontroller P89V51RD2, RS232, MAX 232 IC (For serial communication), microcontroller board, Keil IDE (Fig.1).

Fig. 1: Front view of camera and Back view of camera

Human detection

The video input of the indoor arena or scene was continuously captured by the cameras (left and right camera), then the video obtained from any one of the two cameras is sampled. One of the two cameras was fixed as a Base camera (Fig. 2). This camera was used throughout the human detection process. A reference image was taken. Now each frame was sampled and it was background subtracted with the reference image, from which the silhouette of the objects present in the scene will be acquired. Few sample silhouettes are shown in (Fig. 3).

Now the obtained silhouette was cross checked with those of the silhouettes present in the codebook (a Database which contains number of silhouettes of the human taken in different postures) for calculating match scores. If the match score obtained with any instance in the codebook exceeds a threshold pre-defined value, then obtained silhouette was considered to be humans'. A threshold is set for the number of matches with the silhouettes present in the code book above which obtained silhouette will be regarded as human and we have got positive results for the sample image.

When a human presence in the sampled frame is detected, the Centroid of the human silhouette is taken such that its (X, Y) co-ordinates of the location of Human in the room can be obtained (Fig. 4).

Disparity estimation

Indian J. Innovations Dev., Vol. 1, No. 8 (August 2012) ISSN 2277 - 5390

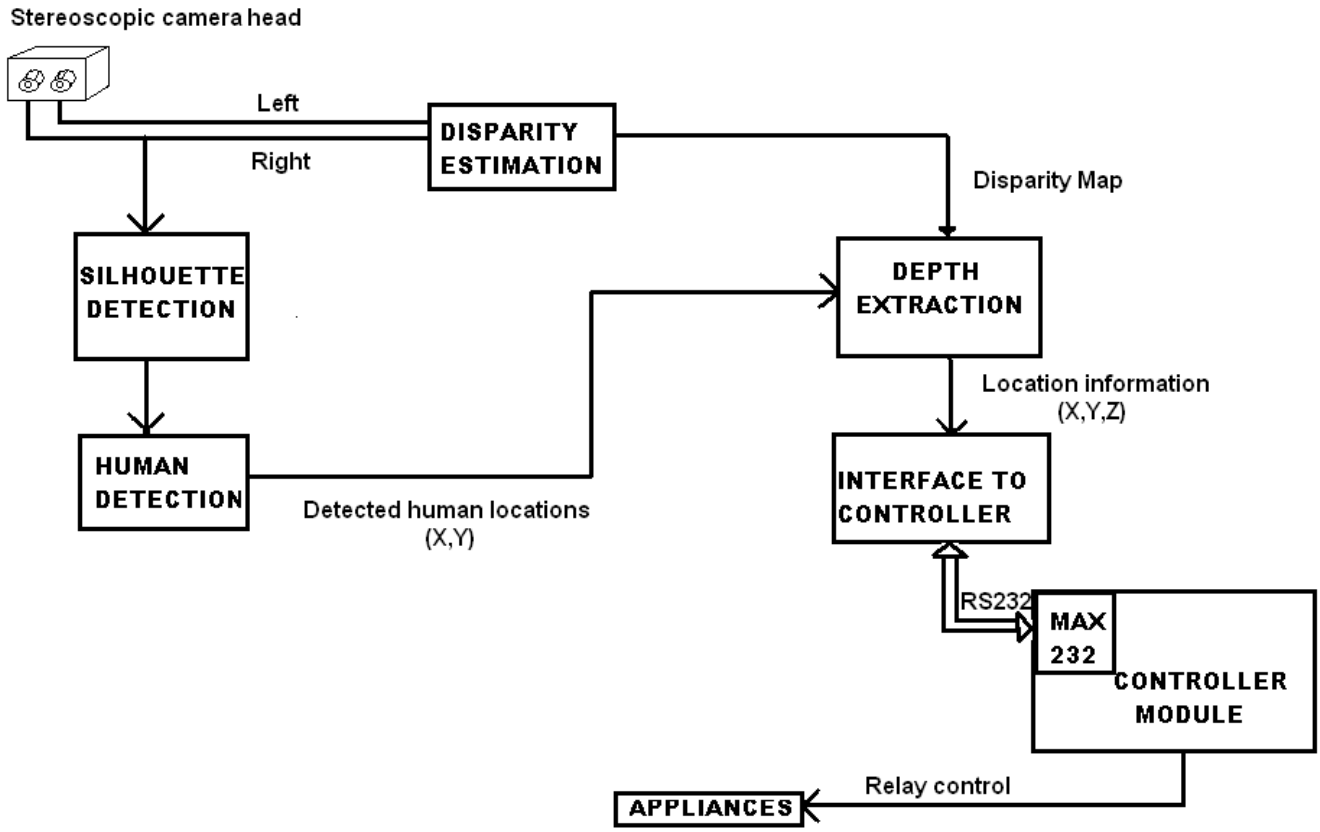

Fig. 2: Overall Block Diagram

Sampled frames of the indoor arena were obtained from the video input, streamed from the right and left cameras of the stereo-pair. Then the two images were used for generating the disparity map between the stereo-pair image below (Fig. 5).

Steps involved in disparity estimation

First take the stereo pair images i.e. Right and left image respectively from two different cameras separated by a fixed baseline distance. Then convert both into intensity images (i.e. RGB to GRAY scale). Consider any one of the two images as reference image. Here let it be Right image. Now a pixel along with its surrounding block from left side image is taken and the block size is an odd number, such that a pixel is chosen along with equal number of pixels on both the side of the centre pixel in both horizontal and vertical direction. Now both the image is checked for shift in horizontal direction, if shift in vertical direction is found, then the program will show an error message. Then both the images are checked for equal number of pixels i.e. same resolution. Now the reference image is scanned within the maximum and minimum disparity range specified to calculate the squared difference between the pixels of the block, which are then added together to get a score and this is called SSD score (i.e. sum of squared difference score). Thus, the SSD method tries to reduce the score and thus the block

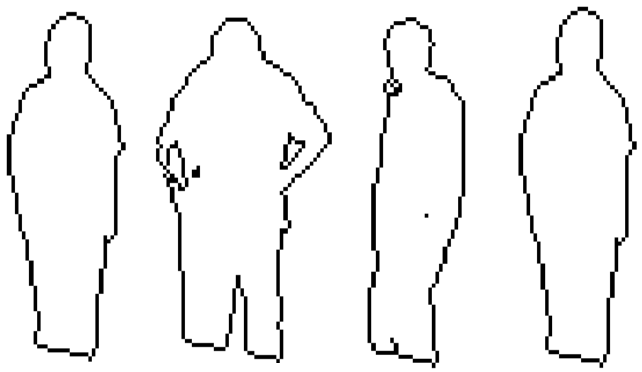

Fig. 3: Human silhouette

Indian J. Innovations Dev., Vol. 1, No. 8 (August 2012) ISSN 2277 - 5390

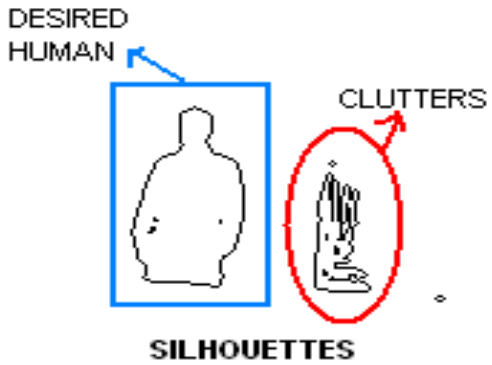

Fig. 4: Silhouettes

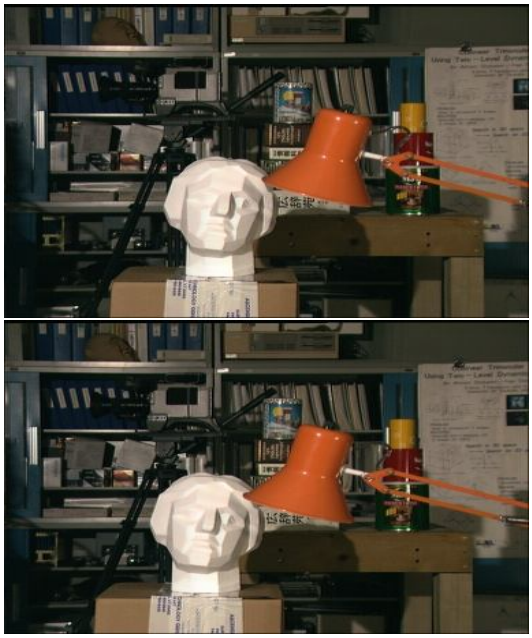

Fig. 5: The left and Right stereo-pair images (ground truth)

with the minimum most score gives the disparity. This procedure is repeated for all the pixels in the image such that, disparity of the entire image is obtained. With this depth map of the image is calculated through which the location (i.e. position) of the human in the room is obtained (Fig. 6).

The disparity was calculated for all the pixels in the image with the method mentioned above. When a human presence in the sampled frame is detected, the foregoing centroid approach was used to obtain the (X, Y) location of the human in the Cartesian space, which in turn was used to obtain the corresponding disparity. The depth information was calculated from this disparity value obtained. Thus the Z co-ordinate of the Cartesian space is obtained. Now the 'z' co-ordinate gives us the 3-dimensional location point of the human in the room can be inferred. The disparity Map obtained for the sample image is shown.

Microcontroller interface

The position obtained from the above phase (i.e. through disparity estimation) is transmitted to Phillips microcontroller P89V51RD2 through RS232 via serial communication (UART).

Basically the room is divided into arrays by calculating disparity at different places in trial and error method. So if the disparity of the human at particular point is obtained, then it will be checked for the limits, so that the array in which the human presents can be found. Since RS–232 is a "complete" standard, it includes more than just specifications on electrical characteristics. The second aspect of operation that is covered by the standard concerns the functional characteristics of the interface. This essentially means that RS–232 has defined the function of the different signals that are used in the interface. These signals are divided into four different categories: common, data, control, and timing. The basic signals used up in this are:

- Transmitted data (TD) ,Received data (RD), Request to send (RTS),Clear to send (CTS), Data terminal ready (DTR),Data carrier detect (DCD)
- The port helps in connecting the controller board to the laptop or system. Figure shows the cable used in connecting and burning the code and also the female part present in the board (Fig. 7).

The serial communication is performed by specifying the port and baud rate of use and burning the needed code to perform our task. The LED's present in the board could be used to display the output. Each LED could be assumed to represent any part of a room that is considered and the presence of human, if detected after the mentioned two processes could be explicitly shown by letting the LED glow i.e. keeping it ON while if there is no human present then the LED will remain OFF and the process will continue by glowing and switching the LED off according to the human presence. In real time application a relay can be operated by

Indian J. Innovations Dev., Vol. 1, No. 8 (August 2012) ISSN 2277 - 5390

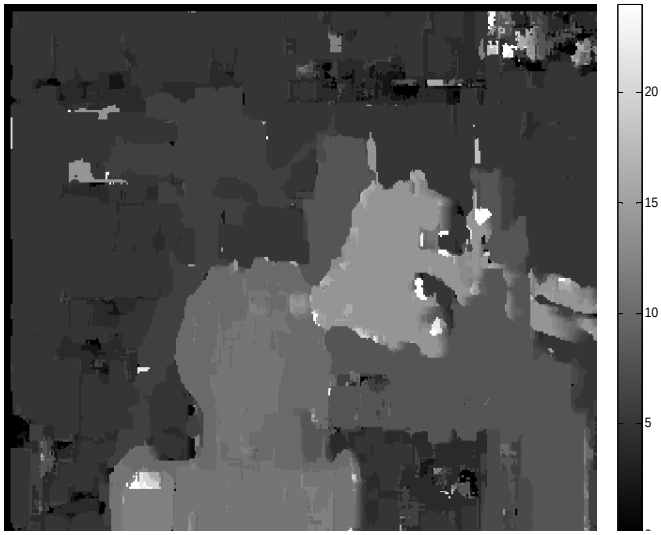

Fig. 6: Their Disparity Map.

Fig. 7: Male jack of DB 9 connector.

Fig.8: Female counterpart

microcontroller output which in turn controls the switches of the concerned area (Fig. 8).

Summary and Conclusion

The very basic idea of power saving with the help of image processing techniques is thus achieved. The process of human detection, though quite difficult to achieve at the very first go, on continuous work and research can be achieved and with the help of disparity estimation the position of human in the concerned area can be obtained. With this the appliance can be controlled effectively. Thus, instead of wasting electricity with the usage of unwanted appliances where the use is not needed, it can be saved to a large extent. Given the power crises around the world, the idea if and when implemented will do wonders with respect to both power saving as well as advancement of technology. This is the first application of its kind where all the above mentioned techniques have been integrated to give fairly optimistic results. As we all know, no work will be flawless. Thus, is the case with ours too.

The little constraints like process time, which is slightly on the higher side (in the order of seconds) could be reduced to a few milliseconds so that things could happen thick and fast as a lay man would want it to. Also efforts could be taken to find out easier methods of human detection such as ratio-proportion or the like which are used for face detection. The above mentioned method gave good results for face recognition where the face is going to be near the camera eye. However, in our case where the human could be at varying distance from the camera, the results were fluctuating. Hence, some other human detection algorithm could be applied to fasten things from what it is now. Stereo images may not always be correctly aligned to allow for quick disparity calculation. Image processing for sure has reached newer heights in its field of application with this work. This gives the hope that the field of work of engineering could be extended to any extent for the betterment of lifestyle. In real time application a relay can be operated by microcontroller output, which in turn controls the power appliances.

References

1. Seyyed Meysam Hosseini, Abbas Nasrabadi, Peyman Nouri and Hasan Farsi (2010) A Novel Human Gait Recognition System. *International Journal of Computer and Electrical Engineering*, Vol. 2(6), 1793-8163.

Indian J. Innovations Dev., Vol. 1, No. 8 (August 2012) ISSN 2277 - 5390

2. Brian De Cann and Arun Ross (2010) Gait Curves for Human Recognition, Backpack Detection and Silhouette Correction in a Nighttime Environment. Proc. of SPIE Conference on Biometric Technology for Human Identification VII, (Orlando, USA), April 2010

3. Murat EK, INC I, Eyup GEDIKLI (2005). Silhouette Based Human Motion Detection and Analysis for Real-Time Automated Video Surveillance. *Turk J Elec Engin*, Vol.13 (2), 199-229.

4. Piotr Skulimowski and Pawel Strumillo (2006) Visual Person Tracking In Sequences Shot From Camera In Motion**.** *Computer Vision and Graphics*. pp 683–688

5. Yigithan Dedeoglu, Ugur Toreyin, B Ugur Gudukbay and Enis Cetin (2006) Silhouette-Based Method for Object Classification and Human Action Recognition in Video**.** LNCS 3979, pp. 64 – 77.

6. Robert T. Collins, Ralph Gross and Jianbo Shi (2002) Silhouette-based Human Identification from Body Shape and Gait. Carnegie Mellon University

7. Shao-Yi Chien, Shu-Han Yu, Li-Fu Ding, Yun-Nien Huang, and Liang-Gee Chen (2003). Fast disparity estimation algorithm for mesh-based stereo image/video compression with two-stage hybrid approach. Visual Communications and Image Processing. Proceedings of SPIE. Vol. 5150, 0277-786X/03.

8. Edgar Arce, Marroquin JL (2007) High-precision stereo disparity estimation using HMMF models. *Image and Vision Computing*, 25, 623–636

9. Patricia Compañ, Rosana Satorre, Ramón Rizo (2003). Disparity estimation in stereoscopic vision by simulated annealing. Group i3a: *Industrial Computing and Artificial Intelligence*

10. Fang Liu. Vasanth Philomin (2008) Disparity Estimation in Stereo Sequences using Scene Flow. Philips Germany

11. Yen San Yong and Hock Woon Hon (2008) Disparity Estimation for Objects of Interest. World Academy of Science, *Engineering and Technology*, 43.

12. Anwar Hasni Abu Hasan, Rostam Affendi Hamzah, Mohd Haffiz Johar (2009). Range Estimation in Disparity Mapping for Navigation of Stereo Vision Autonomous Vehicle Using Curve Fitting Tool. *International Journal of Video & Image Processing and Network Security* IJVIPNS Vol 9(9), 101-105.

13. Zhiguang zhong, Jianqiang Yi, Dongbin Zhao, Yiping Honh (2005). Effective pose estimation from point pairs" Journal Image and Vision Computing archive Vol 23(7), 651-660.

14. Siraj Sabihuddin and James MacLean W (2007). Maximum Likeli hood Stereo Correspondence using Field Programmable Gate Arrays" Proceedings of the 5th International Conference on Computer Vision Systems.

15. Sakuragi K, Kawanaka A (2010) Depth estimation from stereo images using sparsity Signal Processing (ICSP), IEEE 10th International Conference.