

A Low Cost Outdoor Service Robot

Norma Elva Chavez¹, Saiph Savage^{2*} and Rodrigo Savage³

¹Department of Computer Engineering, National Autonomous University of Mexico, Mexico

²Computer Science, West Virginia University, Morgantown, West Virginia, USA

³Computer Engineering, University of Arizona, Tucson, Arizona USA

Abstract

Most service robots are simple and domestic, functioning only within Wi-Fi enabled homes. Service robots that do function outdoors are usually expensive and consequently foreign to the general public. To provide service robots to everyone, everywhere, we present Nokovic, a low cost service robot built with recycled parts based on FPGA and radio frequency control that recognizes outdoor and indoor human activity to engage with humans in any environment. By building off recycled pieces, Nokovic reduces the costs to construct the recycled robot. Nokovic also implements a novel low cost human activity recognition system that further helps to reduce the resources needed for the service robot to function anywhere. We establish our human activity recognition system by using a histogram of oriented gradients to create two shape based object classes: bikers and walkers, which our system distinguishes from. We study how our human activity recognition method compares with traditional activity recognition classifiers. Based on our design and study, we reflect on the future of low cost service robots for outdoor and indoor activities.

Keywords: Service robots; Human recognition module; Nokovic

Introduction

Service robots have become relatively common in the homes of individuals, where they perform services that help people live easier lives [1]. For instance, robots can provide indoor navigation assistance to elders and people with disabilities, allowing them to live normal lives outside of the hospital [2]. Service robots also enable people to clean their houses faster, and have overall better services, such as throwing out the trash [3], despite the increase in service robots, there is a limited number of service robots available to people [4]. The first type of service robot is usually a robot which performs one task, e.g., cleaning floors. The tasks this service robots conduct is generally limited to indoors and has few interactions with humans [5]. The second type of service robot is one that resembles a smartphone with a stand [6]. This type of service robot can move and interface with smart appliances in people's homes but cannot do anything beyond what your phone currently does. As a result, this type of service robot is also typically confined to indoor spaces. The last type is a robot for the sole purpose of entertainment. The latter are mainly machines that are meant to be friendly and interactive. This type of service robot usually acts as a friendly indoor pet which does not function or interact outside a home setting [7]. As we observe, most service robots function primarily within indoor spaces. Services robots that can interact in a wide variety of environments are rare and costly. For instance, Willow Garage offers PR2, a personal service robot with a vast number of sensors to do just about any task in any environment. However, this type of robot also costs over 400,000 dollars.

To enable low cost service robots that can function both indoors and outdoors, we present Nokovic. Nokovic is a service robot that is built with recycled material and based on FPGA and radio frequency control to reduce construction costs. Nokovic integrates detection of outdoor physical human activities to enable the robot to provide outdoor services. The detection of human activity is done with off the shelf cameras and uses a novel low-cost human activity classifier we have proposed, which considers that human activity, especially outdoor activity (humans walking and humans biking), can be recognized by using appearance as the only discriminative factor. Through Nokovic, we showcase how we can build low cost service robots that can function

and engage with people both outdoors and indoors. Figures 1 and 2 showcase Nokovic engaging outdoors and indoors (Box 1).

Nokovic

Nokovic is a service robot that is built from recycled material to reduce its costs. The robot is composed of two main components to enable its functionality: a control module that mobilizes and makes decisions for the robot (controls and decides how the motors will



Figure 1: Photo of Nokovic the service robot engaging outdoors.

***Corresponding author:** Saiph Savage, Computer Science, West Virginia University, Morgantown, West Virginia, USA, Tel: 011525556223067; E-mail: saiph.savage@mail.wvu.edu

Received July 14, 2017; **Accepted** August 04, 2017; **Published** August 14, 2017

Citation: Chavez NE, Savage S, Savage R (2017) A Low Cost Outdoor Service Robot. Adv Robot Autom 6: 168. doi: [10.4172/2168-9695.1000168](https://doi.org/10.4172/2168-9695.1000168)

Copyright: © 2017 Chavez NE, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



Figure 2: Nokovic in action indoors.

Nokovic's Modules

- Service Robot Structure**
made from Recycled Material to Reduce Cost
- Service Robot Control**
Uses FPGA and servo-motors for movement outdoors
- Service Robot Vision**
Recognizes and tracks outdoor human activity to engage with people outdoors

Box 1: Nokovic's modules.

move), and a computer vision module that detects human activity. This recognition is then fed into the control module to enable the robot to make decisions and interact with humans who are moving about in the environment. Nokovic's design is focused on reducing the cost of constructing the service robot, and also on enabling interactions with the robot outdoors.

Nokovic is composed of three modules: structure, control, and vision that enable it to engage with humans in a low cost way within outdoor settings.

Nokovic's structure module

To keep the costs of Nokovic's construction low and to enable a wider range of people to access service robots, we used recycled material to build the structure of Nokovic. To select the recycled material to construct the robot's structure, it is necessary to define the type of operation that the robot will do. For example, if the robot needs to carry items, such as clothing or food, the robot then needs to have a body structure where items can easily be placed and transported. The robot might need to have a place where it can hang things or where boxes can be accommodate with food and water. If the robot is completing tasks that require the robot to move its arms, the robot then needs to have long electric arms. We defined the set of tasks that we

wanted our robot to do, and then used material available in recycle bins to complete the robot's design. For Nokovic's aesthetics, we also used recycled material. In our design we considered it was important for the service robot to have humanistic features, or a "futuristic robot look," similar to the science fiction stories of robots. We considered that by following such aesthetics, people would be able to better identify the service robot and interact with it. We used old cd's to mimic the robot's eyes along with little stones. By using recycled material creatively, we were able to build with low cost the structure of a service robot and established the decorative material that will serve for the service robot's aesthetics.

Nokovic's service robot control

We designed a control module to orchestrate how Nokovic would interact with humans and move about the environment, especially outdoors. For this purpose, Nokovic's control module is composed of two blocks: one external block and one internal block, each with different elements that enable the control and manipulation of the robot. The external block is composed of the motors to move the robot, the FPGA (Spartan-3) that states how the robot should move and what tasks to do, a radio frequency module (RF) that signals to the FPGA what tasks the robot should do, and input and output devices to enable each of these pieces to communicate with each other.

The necessary external elements can be seen in Figure 3.

Figure 3 shows the elements that interact with the control of the service robot.

After designing the external block diagram, we designed the internal block diagram, which aid the VHDL language design for the proper operation.

We can see that one important element to be considered within the circuit that controls the servomotors is a frequency divider, which allows internal management of the external main clock. This external main clock oscillates at a frequency of 50 MHz. Since the robot's digital servomotors are controlled with pulses, whose duration can vary between 600 and 2400 microseconds in high, it is necessary to divide the main external clock of 50 MHz to one of 200 kHz, due to the fact that a clock signal of 200 kHz has a period of 5µs. Thus, we need to divide the movement of each of the servomotors into one that goes from -95° to 95° to cover 360 different positions. This wide range of movements allows Nokovic to better navigate the environment, which is especially important in outdoor settings. Note that the same clock of 200 kHz was used for the diverse elements of the programmed circuit, since in this control we needed counters that carried out delays that lasted seconds and that had different sequences.

We used the Radio Frequency to send the tasks to the FPGA. There are 16 different tasks the robot can receive. We used the popular encoder and decoder HT12E and HT12D together with the transmitter

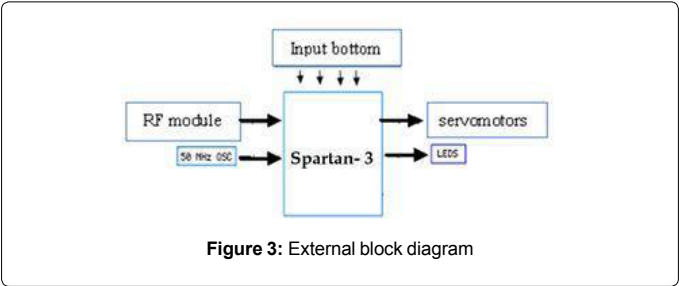


Figure 3: External block diagram

module tws-bs3 and the communication between the different elements. We used the Frequency Range of: 433.92 MHz. Figure 4 shows the connections diagram of the transmitter module.

The RF receiver module is added to the robot, and the algorithm that processes the information received runs on the FPGA, while user controls the RF Transmitter module.

Algorithm for different tasks: The algorithm runs on an FPGA programed in VHDL (VHSIC Hardware Description Language). It is deployed for parallel processing. The FPGA receives four bits from the transmitter module, and these bits allow the robot to select one out of sixteen tasks. The robot is autonomous. There is a predefined path for the robot to transverse. Although the path is defined, the room dimensions and shape are not defined, which means it works in any environment.

Nokovic's service robot vision

This section gives an overview of how Nokovic's framework for recognizing human activity functions. This is summarized in Figure 5.

This module is divided into two different parts, which operate independently, but are subsequently incorporated for the final outcome, in which an instance in the scene is classified and afterwards tracked. Its motion analysis is also obtained, and we are able to predict where the instance will appear in the next frame. This enables our robot to be able to track outdoor human activity in real time.

Offline training of the human recognition module

Nokovic's human activity recognition module is composed of an offline training part whose purpose is to return a classifier that is able to discriminate in a low cost manner human physical activities, especially activities that happen outdoors. Our idea is that by enabling Nokovic to recognize human activities, Nokovic will be able to interact with humans outdoors. In order to learn different human physical activities, we utilized two different learning methods, SVM (Support Vector Machine), and KNN (K-Nearest Neighbor). We compared both methods in a quantitative manner. The training instances that were used

were descriptors of images with labels. In this case, we feed our learning images a series of images of human activities with labels, for instance a photo of a man walking with the label "walking, and for each image calculate its descriptor (a feature vector that characterized the image). These descriptors were computed utilizing two distinct methods, HOG (Histogram of Oriented Gradients) and PHOG (Pyramidal Histogram of Oriented Gradients). We worked with these descriptors because we wanted to explore whether we could characterize human activity with low cost recognition methods, in this case by simply considering local appearance and shapes that exist in an image and that are described by the distribution of intensity gradients or edge directions [8]. We also carried out quantitative recognition comparisons between these descriptors and classifiers. By using HOG and PHOG for obtaining the descriptor, we were able to explore whether we could create a classifier of human activity that used only human appearance for its classification, and hence was more low cost than traditional methods.

Our classifier was trained and tested using three different data sets. We aimed for these datasets to have a variety of human physical activity so that our robot could potentially recognize the things humans were doing outdoors to better engage with them. The first dataset we collected is a very well supported MIT pedestrian database that consists of 709 images of pedestrians in cities. Because this dataset contains a relatively limited range of poses and primarily has only walking activity, we sought to incorporate other image databases. One of these resulting databases was the INRIA dataset, which contained over 1805 images of people in practically any orientation and against a wide variety of backgrounds including crowds. The only problem we encountered with this database was that although it did hold a small subset in which humans were doing other activities (approximately 40 images were of humans biking) the number of these images was not sufficient to train and test our classifier properly. Therefore we produced a new data set from a varied set of personal pictures as well as images taken from Google which contained photos of humans doing diverse physical activities, especially biking. We wanted first to evaluate whether we could devise a low cost classifier that could detect at least two different outdoor human activities, in this case: walking and biking.

Offline training methodology

Our goal was to probe the effectiveness of our classifier, especially to classify human activity. For this purpose, we first divided our dataset into three different classes: pedestrians, bikers, and person-free photos (we wanted Nokovic to be able to also recognize when humans were not in the picture to interact differently). Each of the classes was assigned a particular label. In this case "Walking", "Biking" or "Background" (a photo without a human). From these images, we selected for each class 200 for training and the rest for testing. (Approximately 1024 images were used for testing each class) In the learning period we computed the descriptor of every image in the training pool. Each descriptor with its corresponding label represented a learning instance. For obtaining the descriptor, we implemented in Open CV-C++ the HOG [8] and the PHOG [9]. The design parameters that were taken in HOG were the number of bins that the histogram has, the number of pixels per cell, and the number of cells per block. In our testing, we varied these parameters to observe what would result in the lowest classification error. On the other hand, the design parameter that was taken in PHOG was the number of bins and the level of the pyramid. In our testing we also modified these values for the purpose of observing which resulted in the lowest error rate.

Once we had the descriptors from all of the learning instances, we trained a SVM using the machine learning library from Open CV

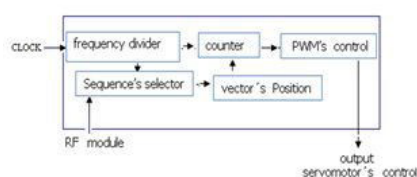


Figure 4: Internal block diagram of the Spartan-3 Starter Kit Board, which we use to control our service robot.

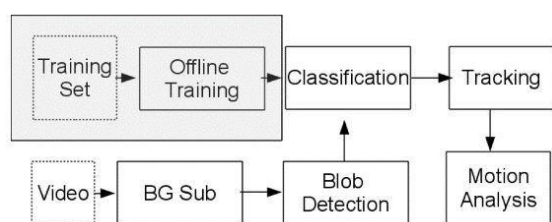


Figure 5: Overall workflow of Nokovic's framework to recognize human activity. The enclosed part is the one that is executed offline, the remaining part flows on the fly.

2.1. The SVM from Open CV 2.1 has many parameters that can be modified. We only experimented with the type of SVM that we used, and we selected between training with a linear SVM or a Radial SVM. The other learning method with which we experimented with was K Nearest Neighbor. In our results we compare the performance of these three learning methods. We tested the effects that diverse parameters in the descriptor and in the learning method had, with the purpose of finding the most robust classifier for our problem. At the end of this step we counted with an activity recognition classifier that could differentiate between different human activities and also detect when a human was present or not in the scene. We present the results of this analysis in our evaluation section [10-13].

Blob detection in the human recognition module

Once we had our classifier which could classify the type of activities that humans were doing, we needed to then develop parts that could in real time detect and follow humans in a scene in order to later establish what they were doing. For this purpose, we developed the blob detection piece that focuses on detecting humans in real time. Nokovic has a camera attached to its head that streams live video of the scene. In this video, we first fulfill background subtraction, and in this form, focus only on the moving objects of the video. After this, we perform a binarization of the image and run a median filter to reduce the noise in the video. We use especially the median filter because it is known for reducing the salt and pepper noise that occurs after thresholding, especially in videos that are unstable like the one we mounted on Nokovic. After we have a more clean video image, we perform connected component analysis on the corresponding binarized filtered image and are able to obtain the blobs (i.e., the parts of the video where humans are likely to be present). Finally, we compute the bounding boxes containing the white pixel portions of the image. From this bounding box we take into consideration its aspect ratio, with the purpose of discarding possible false positives. We consider that human pedestrians and bikers present a certain relationship in their shape and thus when the expected aspect ratio is not met, the blob is discarded. At the end of this step we return a list of blobs that were detected in a particular time frame [14].

Classification and motion analysis in the human recognition module

We combine the outcomes from Blob Detection and Offline training to obtain a real time classifier that can continuously track humans, and detect the activities they are doing in a low cost form. This enables Nokovic to be able to interact with humans outdoors, as the robot knows what humans are up to. For this purpose, we feed into our classifier the blobs that were detected in the previous step. Each blob has been previously scaled to a size of 64×128 , and the classifier just focuses on predicting its corresponding label.

After the activity has been recognized, we begin the tracking of the object, which is done through the Pyramidal Lucas-Kanade Opencv [10] implementation. The features which we track are the key points from the Harris Corner Detector; these features were selected since Shi-Tomasi corners provided a non-stable tracking. From each tracked blob, we compute its trajectory and also obtain its average speed. The trajectory is represented by a two-dimensional N-tuple corresponding to the x-axes and y-axes projections of the blob's centroid location at each instant of time:

$r[k]=x[k]; y[k]; k=1; \dots; N$. At the end of this step, Nokovic is able to detect the activities that humans are doing, and follow them.

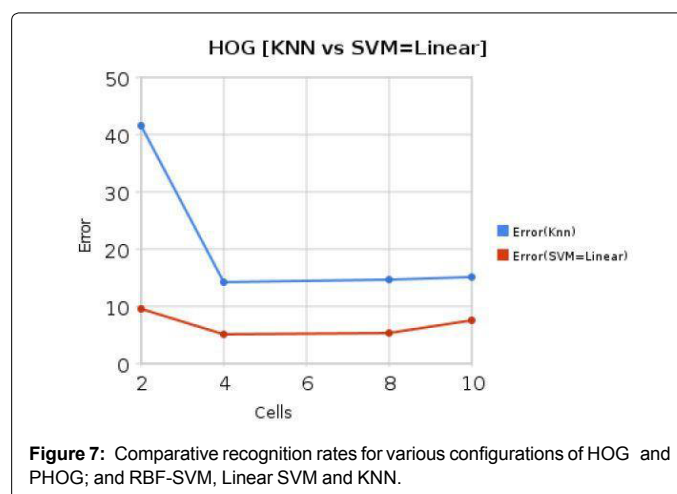
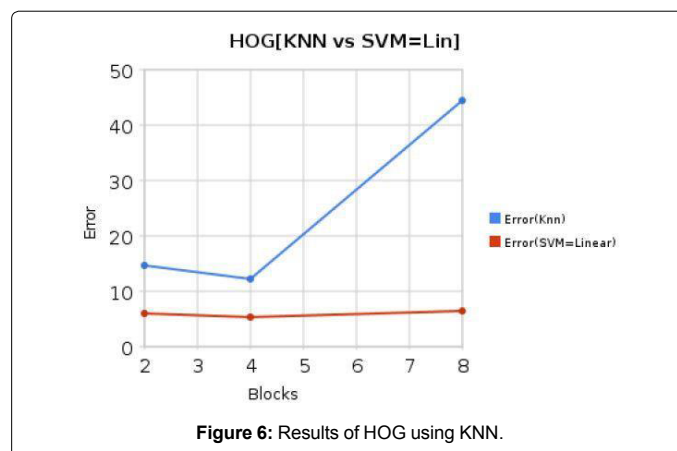
Experimental Results

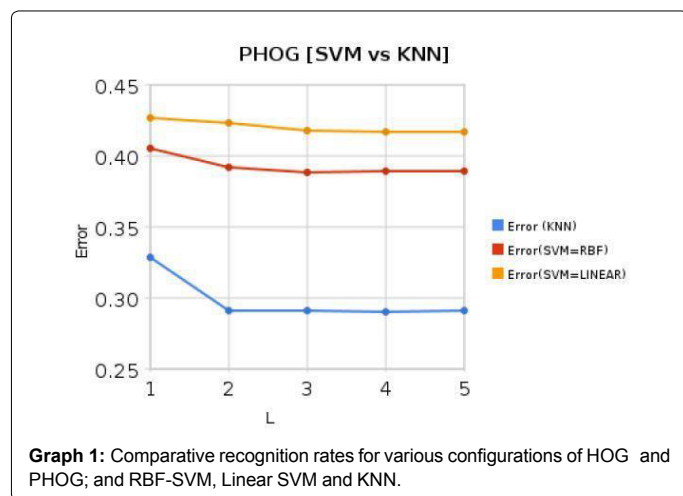
We were at first hand interested in finding which of the classifiers was the most robust for discriminating between different human activities. For this purpose, we explored and altered the design parameters of each different classifier. We present the results of these modifications, and they were altered, and the resulting effects of these modifications are shown in the graphs we plotted.

We initially compared RBF-SVM with Linear-SVM. In this case, the descriptors of the images were computed using HOG with the parameters, i.e., number of cells per block, number of pixels per cell, number of bins etc., that the paper [8] mentioned were the best for pedestrian detection. From this experimentation we found that the linear SVM was the one that resulted in the lowest error rate in human activity recognition. We found that Linear SVM presented an overall error rate of 6:06% while Radial SVM presented an overall error rate of 14:18%, as shown in Figure 5.

We also compared linear SVM vs. KNN and vs. Radial SVM using PHOG descriptors. In this experiment we found that SVM outperforms both of the other learning methods, although the difference between SVM and KNN may not be that wide, as it was in first instance expected Figures 6 and 7 (Graph 1).

Now in terms of the descriptors, there were also design parameters that were modified. The results of these modifications are presented in the graphs 1. From our experimentation we saw that when HOG





descriptors were utilized, the performance of all the classifiers improved significantly compared to when PHOG descriptors were used. In HOG descriptors, we found that the optimal number of pixels per cell was 4, and the optimal number of cells per block was two.

We also performed experimentations on video surveillance that Nokovic took while navigating the campus of the national autonomous university of Mexico (UNAM). We used Nokovic's activity recognition module to recognize two different human activities: Biking and Walking. The ground truth was established manually. We observed that we were able to detect different human activities in the university. Our next steps will be in defining how Nokovic will interact with humans after it has detected that they are performing a certain activity.

Conclusion

In this work we presented Nokovic as a low cost service robot that functions both outdoors and indoors. With Nokovic we were able to showcase how we can build service robots out of recycled material, and which can detect outdoor human activity to engage with humans outside the home. Our work explored different configurations among descriptors (HOG and PHOG) and classifiers (SVM and K-NN) with the aim to effectively recognize different human activities, namely Biking and Walking, based on appearance and thus reducing the computational costs. Although our results suggest that to obtain a more accurate classification of human activities it is necessary to take into account additional descriptors, the results suggest that we can use appearance to start to recognize human activity in a low cost way. In future work it might be interesting to combine appearance along with other cues such as motion and to observe how this improves the prediction of human activity. Furthermore, for Blob detection it has been proven that morphological operations can improve this task, so it would be interesting to integrate this into the system to obtain a better performance. Shadows were not addressed in this implementation, and it caused problems to accurately track some pedestrians, especially within the real world setting. Therefore, shadows must be addressed to ensure a better outcome.

In future work, we plan to run user studies with our robot to

further understand how people engage with service robots that are low cost and that can potentially follow them everywhere (both outside and inside). We will also explore different services for the robot to provide and inspect how people perceive the quality of the work that the robot provides in different settings. We believe that by opening up the design space to create low cost service robots, we can empower a larger number of people to have access to robots. We are interested in studying how this new form of interactions plays out, especially how people perceive them. In the future, we will also explore ways to empower people to design and build their own low cost service robots. We have been able to devise an approach that facilitates the creation of service robots anywhere. We would like to explore how we can now also enable people to design the type of outdoor service robots they need.

References

- Jodi F, DiSalvo C (2006) Service robots in the domestic environment: a study of the roomba vacuum in the home. *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM.
- Ho Seok Ahn, Min Ho Lee, Broadbent E, Macdonald B (2017) Gathering Healthcare Service Robot Requirements from Young People's Perceptions of an Older Care Robot. *IEEE International Conference on Robotic Computing (IRC)*.
- Bharadwaj MRK, Menassa CC, Kamat VR (2017) Task Allocation and Route Planning for Robotic Service Networks in Indoor Building Environments. *Journal of Computing in Civil Engineering* 31: 04017038.
- Ivanov SH, Webster C, Berezina K (2017) Adoption of robots and service automation by tourism and hospitality companies. *INVTUR Conference*, Aveiro, Portugal.
- Koji K, Nishio S, Hagita N, Sato M (2012) Cloud networked robotics. *IEEE Network*.
- Sanjay D, Chinnaiiah MC, Sai Kiran Ch, Sai Priyanka B, Punna Rao P (2016) An FPGA based service Robot for floor cleaning with autonomous navigation. *International Conference on Research Advances in Integrated Navigation Systems (RAINS)*, IEEE.
- Shin N, Ohtani N (2016) A pet-type robot AIBO-assisted therapy as a day care program for chronic schizophrenia patients: A pilot study therapy. *Australian Medical Journal* 9: 244-248.
- Gentili GB, Tesi V, Linari M, Marsili M (2002) A versatile microwave plethysmograph for the monitoring of physiological parameters. *IEEE Trans Biomed Eng* 49: 1204-1210.
- Beebe DJ (1993) Signal conversion. In: *Biomedical Digital Signal Processing*, Tompkins WJ, Editor Englewood Cliffs, NJ: Prentice-Hall 3: 61-74.
- Akay M (1998) *Time Frequency and Wavelets in Biomedical Signal Processing*. Piscataway, NJ: IEEE Press, pp: 123-135.
- Miller EH (2017) A note on reflector arrays. *IEEE Trans. Antennas Propagat.*, in press.
- Menendez T, Achenbach S, Moshage W, Flug M, Beinder E, et al. (1998) Prenatal recording of fetal heart action with magnetocardiography. *Z Kardiol* 87: 111-118.
- Monzon JE (2000) The cultural approach to telemedicine in Latin American homes. *Proceedings 2000 IEEE EMBS International Conference on Information Technology Applications in Biomedicine. ITAB-ITIS 2000. Joint Meeting Third IEEE EMBS International Conference on Information Technology*, Arlington, VA, pp: 50-53.
- Saunders FA (1973) Electrotactile sensory aids for the handicapped. Presented at the 4th Annu. Meeting Biomedical Engineering Society, Los Angeles, CA.