Controlled of Mobile Robots by Using Speech Recognition

Ahmed Q. AL-Thahab
Babylon University,College of Engineering

Abstract
This paper presents a proposed technique of speech recognition system and it applies to voice control of electromechanical appliance, especially voice controlled mobile robots or intelligent wheelchair for handicapped people. Our aim is to interact with the robot using natural and direct communication techniques. The aim of this paper is that how the voice can be processed to obtain proper and safe wheelchair movement by high recognition rate. In order to make voice an efficient communication tool between human and robots, high speech recognition rate must be achieved. But one hundred percent speech recognition rate under a general environment is almost difficult to achieve. In this paper, proposed technique called (Multiregdilet transform) is used for isolated words recognition. Finally use the outputs of neural network (NNT) to control the wheelchair through computer note books and special interface hardware. A successful recognition rate of 98% was achieved.
Keywords: Artificial Neural Network, Multiridgelet Transform, Multiwavelet Transform, and Interfacing Circuit.

الخلاصة

قدم هذا البحث تقنية مقترحة في الأنظمة التي تميز الكلام والتي وضعت لاستخدامها في السيطرة على المعدات الالكتروميكانيكه (electromechanical)، خصوصاً السيطرة على الإنسان الآلي المتحرك أو الكرسي المتحرك الذكي للأشخاص المعوقين. هدفنا يكون التفاعل مع الإنسان الآلي باستخدام تقنية الاتصال الطبيعية أو المباشرة. الهدف من هذا البحث هو كيف معالجة الصوت للحصول على حركة صحيحة وأمينة عن طريق نسبة تمييز عالية. وكذالك لكي نجعل الكلام أداة الاتصال الكفوءة بين البشر والإنسان الآلي، يجب تحقيق معدل عالي لتمييز الكلام. لكن مئة بالمائة من نسبة تمييز الكلام تحت ظروف جوية اعتيادية هي كذالك من الصعوبة تحقيقها. في هذا البحث، التقنية المقترحة المسماة تحويل متعدد رجلت استخدمت لتمييز الكلام المنفصل. أخيراً استخدام الخارج من الشبكة العصبية للسيطرة على الكرسي المتحرك عن طريق حاسبة كتب ودائرة عملية خاصة متداخلة بين الحاسبة والكرسي المتحرك. نجاح نسبة التمييز هي ٩٨% قد حققت.

1. Introduction

Since human usually communicates each other by voices, it is very convenient if voice is used to command robots. A wheelchair is an important vehicle for the persons physically handicapped. However, for the injuries who suffer from spasms and paralysis of extremities, the joystick is a useless device as a manipulating tool [Komiya, *et al.*, 2000]. The recent developments in speech technology and biomedical engineering world have diverted the attention of researchers and technocrats to concentrate more towards the design and development of simple, cost effective and technically viable solutions for the welfare and rehabilitation of a large section of disabled community [Lim, *et al.*, 1997].

One method is to command the wheelchair by the voice through special interface, which plays role of master control circuit for the motors of wheelchair. In case of voice control there is more difficult situation because the control circuit might generate recognition error. The most dangerous error for wheelchair control is substitution error, which means that recognized command is interpreted as opposite command. For example, "left" is interpreted as "right". Situation described above are very probable in the polish language. A word meaning "left" and "right" has very high acoustic similarity [Sajkowski, 2002].

Robotic arms fitted with some type of gripper, which can be used to help people eat, assist with personal hygiene, fetch items in a home or office environment, and open door knobs. The arms can be mounted on wheelchairs, attached to mobile robots, on a mobile base, or fixed to one location as part of a workstation. An overview of rehabilitation research investigating robotic arms and systems can be found in [Mahoney, 1997]. Arms mounted on wheelchairs must not interfere with normal use of the wheelchair by increasing its size too much or causing the chair's balance to become unstable [Yanco, 2000].

2. Related Work

[Moon, *et al*., (2003)] proposes an intelligent robotic wheelchair with user-friendly human-computer interface (HCI) based on electromyogram (EMG) signal, face directional gesture, and voice. The user's intention is transferred to the wheelchair via the (HCI), and then the wheelchair is controlled to the intended direction. Additionally, the wheelchair can detect and avoid obstacles autonomously wing sonar sensors. By combining HCI into the autonomous functions, it performs safe and reliable motions while considering the user's intention.

The method presented by [Rockland, et al, 1998] was designed to develop a feasibility model for activating a wheelchair using a low-cost speech recognition system. A microcontroller was programmed to provide user control over each command, as well as to prevent voice commands from being issued accidentally. It is a speaker dependent system that can be trained by the individual who would be using the system, and could theoretically attain a better than 95% accuracy. [Sundeep, et al, 2000] they presented a voice control through the feature based, language independent but speaker dependent, isolated word recognition system (IWRS) that uses discrete time warping (DTW) technique for matching the reference and spoken templates. A unique code corresponding to each recognized command is transmitted on a parallel port from the IWRS to the motor controller board that uses 80KC 196 micro-controllers. [Simpson, et al, 2002] they proposed to utilize voice control in combination with the navigation assistance provided by "smart wheelchairs," which use sensors to identify and avoid obstacles in the wheelchair's path. They were describes an experimental result that compares the performance of able-bodied subjects using voice control to operate a power wheelchair both with and without navigation assistance.

[Valin, et al, 2007] they described a system that gives a mobile robot the ability to perform automatic speech recognition with simultaneous speakers. A microphone array is used along with a real-time implementation of geometric source separation (GSS) and a postfilter that gives a further reduction of interference from other sources. The postfilter is also used to estimate the reliability of spectral features and compute a missing feature mask. The system was evaluated on a 200-word vocabulary at different azimuths between sources. [Hrnčár, 2007] describes the Ella Voice application, which is the user-dependant, isolated voice command recognition tool. It was created in MATLAB, based on dynamic programming and it could serve for the purpose of mobile robots control. He deals with the application of selected techniques like cross-words reference template creation or endpoints detection. [Takiguchi, et al, 2008] they proposed speech recognition is one of the most effective communication tools when it comes to a hands-free (human-robot) interface. They describe a new mobile robot with hands-free speech recognition. For a hands-free speech interface, it is important to detect commands for a robot in spontaneous utterances. The system can understand whether user's utterances are commands for the robot or not, where commands are discriminated from human-human conversations by acoustic features. Then the robot can move according to the user's voice (command). Recognition rates for user's request were 89.93%.

From the previous work, it seems that all automatic speech recognition ASR used in wheelchair have low recognition rate less 95% than (speaker dependent) and navigation system depend on additional sensor like IR, ultrasonic, camera….etc.

In this work voice is completely control wheelchair with recognition rate 98% for speaker independent and a paper titled "Controlled Mobile Robots by Using Speech Recognition" published in, and from this rate I conclude that this method is better than the previous methods.

3. System Design

The following 5 voice commands have been identified for various operation of the wheelchair FORWARED, REVERSE, LIFT, RIGHT, and STOP. Chair starts moving in corresponding direction on uttering the command forward in forward direction and stop if the command is stop and so on.

3-1 Data Base of Speech

Every speaker recognition system depends mainly on the data input. The data that used in the system is speech. The speech uttered by using 15 speakers, 8 males and 7 females, 10 of them used for training purpose (5 males, and 5 females) and each speaker utter the same word 5 times. The following 5 voice commands has been selected for various processes of the wheelchair (Forward, Reverse, Left, Right, and Stop). The total numbers of uttered data used for training is 250. The remaining speakers (3 males, and 2 females) are used for tested purpose, and each speaker uttered the same word 2 times, then the total number of utterance is 50 that used for testing purpose.

3-2 Multirighelet Transform

To improve the performance and to overcome the weakness points of the Ridgelet transform, a technique named the Multiridgelet transform proposed. The main idea of the Ridgelet transform is to map a line sampling scheme into a point sampling scheme using the Radon transform, then the Wavelet transform can be used to handle effectively the point sampling scheme in the Radon domain [Minh, *et al*., 2003].While the main idea of Multiridgelet transform depends on the Ridgelet transform with changing the second part of this transform with Multiwavelet transform to improve the performance and out put quality of the Ridgelet transform.

In fact, the Multiridgelet transform leads to a large family of orthonormal and directional bases for digital images, including adaptive schemas. However, the Multiridgelet transform overcome the weakness point of the wavelet and Ridgelet transforms in higher dimensions, since the wavelet transform in two dimensions are obtained by a tensor-product of one dimensional wavelets and they are thus good at isolating the discontinuity across an edge, but we will not see the smoothness along edge. The geometrical structure of the Multiridgelet transform consists of two fundamentals parts, these parts are:

a- The Radon Transform.

b- The One Dimension Multiwavelet Transform.

3-3 The Radon Transform

The Radon transform is defined as summations of image pixels over a certain set of “lines”. The geometrical structures of the Radon transform consist of multiple parts of the sequence jobs. Radon transform provides a mean for determining inner structure of an object. It allows us to analyze signal in detail by means of transforming the original signals from the spatial domain into projection space [Li, *et al*., 2003].

Radon transform (RT) appears to be a good candidate. It converts original image into a new image space with parameters υ and t. Each point in this new space accumulates all information corresponding to a line in the original image with angle υ

and radius t. Thus, when radon transform localizes near an angle υ_o and around a slice t_o a local maximum will results original image that has a line in position (t_o, υ_o). This is the kind of transform we are looking for [Terrades, *et al*., 2003].

4. Neural Network

Artificial Neural Networks (ANN) refers to the computing systems whose central theme is borrowed from the analogy of 'biological neural networks'. Many tasks involving intelligence or pattern recognition are extremely difficult to automate [Ram Kumar, *et al*., 2005].

4-1 The Model of Neural Network

We used random numbers around zero to initialize weights and biases in the network. The training process requires a set of proper inputs and targets as outputs. During training, the weights and biases of the network are iteratively adjusted to minimize the network performance function. The default performance function for feed forward networks is mean square errors, the average squared errors between the network outputs and the target output [Hosseini, *et al*., 1996].

4-2 Back Propagation Training Algorithm:

The back propagation is designed to minimize the mean square error between the actual output of multilayer feed-forward Perceptron and the desired output [Zurada, 1996]. Figure (1) shows the basic two-layer network:

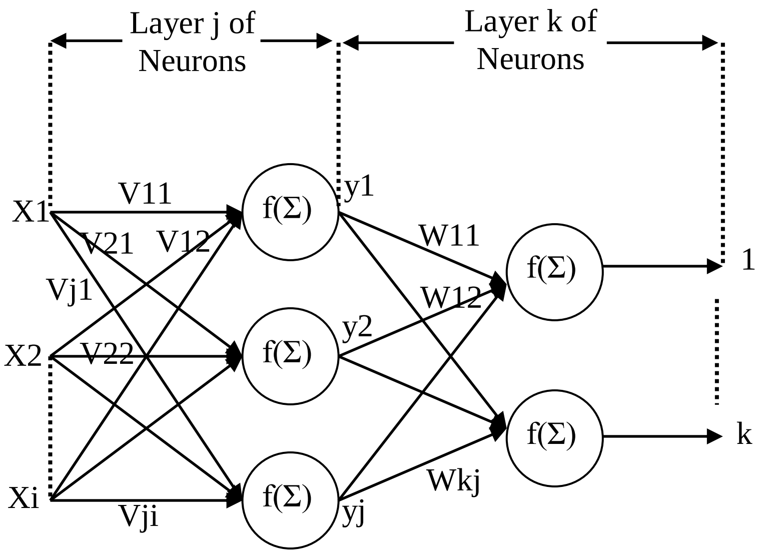

Figure (1), The Back Propagation Network

Summary of the Back-Propagation Training Algorithm (BPTA):

Step1: $\eta > 0$, $E_{max.}$ chosen.

Weights W and V are initialized at small random alues; W is (K×J), V is (J×I).

Step 2: Training step starts here, input is presented and the layer's output computed [f(net)]

$$f(net) = \frac{2}{1+\exp(-\lambda net)} - 1 \qquad (1)$$

$$y_j \leftarrow f(v_j^t x), \text{ for } j = 1,2,3,...,J \qquad (2)$$

Where v_j, a column vector, is the j'th row of V, and

$$o_k \leftarrow f(w_k^t y), \qquad \text{for } k = 1,2,3,...,K \qquad (3)$$

Where w_k, a column vector, is the k'th row of W.

Step3: Error value is computed:

$$E \leftarrow \frac{1}{2}(d_k - o_k)^2 + E, \qquad \text{for } k = 1,2,3,...,K \tag{4}$$

Step 4: Error signal vector δ_o and δ_y of both layers are computed. Vector δ_o is $(K\times1)$, δ_y is $(J\times1)$. The errors signal terms of the output layer in this step is:

$$\delta_{ok} = \frac{1}{2}(d_k - o_k)(1 - o_k^2), \qquad \text{for } k = 1,2,3,...,K \tag{5}$$

The error signal term of the hidden layer in this step is

$$\delta_{yj} = \frac{1}{2}(1 - y_j^2)\sum_{k=1}^{K} \delta_{ok} w_{kj}, \qquad \text{for } j = 1,2,3,...,J \tag{6}$$

Step 5: output layer weights are adjusted:

$$w_{kj} \leftarrow w_{kj} + \eta\delta_{ok} y_j, \qquad \text{for } k = 1,2,3,...,K \text{ and } j = 1,2,3,...,J \tag{7}$$

Step 6: Hidden Layer weights are adjusted:

$$v_{ji} \leftarrow v_{ji} + \eta\delta_{yj} x_i, \qquad \text{for } j = 1,2,3,...,J \text{ and } i = 1,2,3,...,I \tag{8}$$

Step 7: If more patterns are presented repeated by go to step 2 otherwise go to step 8.

Step 8: The training cycle is completed

For $E<E_{max}$ terminate the training session.

If $E>E_{max}$ then $E \longleftarrow 0$, and initiate the new training cycle by going to step 2.

5. General Procedure of Proposed Systems

This paper contain two part, part one contains the theoretical work (simulation in computer with aid of matlab 7), and the second one puts interface between computer and connected to wheelchair. The first part contains three steps for implementation:

a. Preprocessing.
b. Feature Extraction.
c. Classification.

The following section gives the detail of each step:

5-1 The Preprocessing: In this section, the isolated spoken word is segmented into frames of equal length of (128 samples). Next the result frames of each word is converted into single matrix (2- dimensional), and this matrix must be power of two. So the proposed length for all word is 16348 (one dimensional), and this length is power of two and can divided into matrix have dimension (128×128 , and this is 2- dimensional and power of two matrix).

5-2 Feature extraction: the following algorithm was used for the computation of 2-D discrete Multiridgelet transform on Multiwavelet coefficient matrix using GHM four multifilter and using an over sampling scheme of preprocessing (repeated row preprocessing). It contains four fundamental part, these are applied to 2-D signal (word) to get best feature extraction:

1. Input word, and check its dimension.
2. Apply 2-D FFT for the resizing signal (2-D word), its method convert the matrix from Cartesian to polar.
3. Apply Radon transforms to 2-D FFT coefficient.

4. Finally apply1-D DMWT to the radon transform coefficient.

The final coefficient gets the Multiridgelet coefficient. The procedure were applied to 2-D signal to get finally best feature extraction.

5-3 Classification: This step begins when getting on 2-D discrete Multiridgelet transform coefficient. The coefficient splitter into two parts, the first part used as a reference data, and the second one used as tested or classified data. The strong method that can be recognized signal simply is neural network that use an algorithm of back propagation training algorithm as a classifier after training the reference data (coefficient) resulting from 2-D discrete Multiridgelet transform.

Because of the input nodes of NNT set as a vector input, then the output nodes also must be set as a vector (1-D) and its value depend on the desired signal for each word. For the same word, the desired is the same but for other words will be different (i.e. the desired signal differs from word to other).

5-4 General Algorithm for Computing Multiridgelet Transform

To compute Multiridgelet transform, the next steps should be followed:

a. Resizing: Here it is necessary to check for speech dimensions, speech matrix should be a square matrix, N*N matrix, where N must be the power of two.
b. Go to 2_D FFT: Computing two dimensions Fast Fourier Transform.
c. Save DC component: Store DC component and replace its position with zero.
d. Computing the best sequence of directions.
e. Compute the Fourier slices: Finding the Fourier slices that must be taken.
f. Rearranging the coefficients matrix: Sorting the result matrix which is getting from 2_D FFT according to the new positions of computing the best sequence of directions.
g. Back to the spatial domain: Applying 1_D IFFT to each row. The result is (Radon Transform).
h. Applying 1_D DMWT: Here it is required to apply the one dimension discrete Multiwavelet transforms (1_D DMWT) for each row. The result matrix is Multiridgelet Transform coefficients.

5-5 Computation FDMWT for 1-D Signal

By using an over-sampled scheme of preprocessing (repeated row), the discrete multiwavelet transform (DMWT) matrix is doubled in dimension compared with that of the input, which should be a square matrix NxN where N must be power of two. Transformation matrix dimensions equal input signal dimensions after preprocessing. To compute a single-level 1-D discrete multiwavelet transform, the next steps should be followed:

1. Checking input dimensions: Input vector should be of length N, where N must be power of two.
2. Constructing a transformation matrix, W, using GHM low and high pass filters matrices given below:

$$H_0=\begin{bmatrix}\frac{3}{5\sqrt{2}} & \frac{4}{5}\\ -\frac{1}{20} & -\frac{3}{10\sqrt{2}}\end{bmatrix},\ H_1=\begin{bmatrix}\frac{3}{5\sqrt{2}} & 0\\ \frac{9}{20} & \frac{1}{\sqrt{2}}\end{bmatrix},\ H_2=\begin{bmatrix}0 & 0\\ \frac{9}{20} & -\frac{3}{10\sqrt{2}}\end{bmatrix},\ H_3=\begin{bmatrix}0 & 0\\ -\frac{1}{20} & 0\end{bmatrix} \quad (9)$$

$$G_0=\begin{bmatrix}-\frac{1}{20} & -\frac{3}{10\sqrt{2}}\\ \frac{1}{10\sqrt{2}} & \frac{3}{10}\end{bmatrix},G_1=\begin{bmatrix}\frac{9}{20} & -\frac{1}{\sqrt{2}}\\ -\frac{9}{10\sqrt{2}} & 0\end{bmatrix},\ G_2=\begin{bmatrix}\frac{9}{20} & -\frac{3}{10\sqrt{2}}\\ \frac{9}{10\sqrt{2}} & -\frac{3}{10}\end{bmatrix},G_3=\begin{bmatrix}-\frac{1}{20} & 0\\ -\frac{1}{10\sqrt{2}} & 0\end{bmatrix} \quad (10)$$

The transformation matrix can be written as follow:

$$W=\begin{bmatrix} H_0 & H_1 & H_2 & H_3 & 0 & 0 & \dots \\ G_0 & G_1 & G_2 & G_3 & 0 & 0 & \dots \\ 0 & 0 & H_0 & H_1 & H_2 & H_3 & \dots \\ 0 & 0 & G_0 & G_1 & G_2 & G_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (11)$$

After substituting GHM matrix filter coefficients values, a 2Nx2N transformation matrix results.

3. Preprocessing the input signal by repeating the input stream with the same stream multiplied by a constant α, for GHM system functions $\alpha = 1/\sqrt{2}$.
4. Transformations of input vector which can be done as follows:
   a. Apply matrix multiplication to the $2N\times2N$ constructed transformation matrix by the $2N\times1$ preprocessed input vector.
   b. Permute the resulting $2N\times1$ matrix rows by arranging the row pairs 1,2 and 5,6 …, $2N-3$,$2N-2$ after each other at the upper half of the resulting matrix rows, then the row pairs 3,4 and 7,8,…, $2N-1$,$2N$ below them at the next lower half.

Finally, a $2N\times1$ DMWT matrix results from the $N\times1$ original matrix using repeated row.

5-6 Results of Neural Network Training

A six-layer feed-forward network with 512 ‘*logsig*’ neurons in the first input layer, the first hidden layer is 256 ‘*logsig*’ neurons, the second hidden layer is 128 ‘*tansig*’ neurons, the third hidden layer is 64 ‘*tansig*’ neurons, the fourth hidden layer is 32 ‘*tansig*’ neurons and 5 ‘*tansig*’ neurons in the output layer corresponding with five signals that used to start wheelchair. The NN has 16384 inputs for each isolated word command extracted by 2-D Multiridgelet transform.

The resilient back propagation method is used, for all speech command, 15 speakers (8 males and 7 females). For NNT training purpose 10 speakers are used (5 males, and 5 females). Each speaker utters the same word 5 times in different environment for each utterance. Then the number of utterances for each word is 50 times, and then the total number of utterances for all words is 250 that are used in NNT training purpose. The remaining speakers are 5 (3 males, and 2 females), that are used for NNT testing purpose. Each speaker utters the same word 2 times in different environment for each utterance. Then the number of utter of each word is 10 times, and then the total number of utterance for all word is 50 utterances that are used in NNT testing purpose. The training process for five control outputs can be shown in shown in Figure (7) and the training results in table (4).

From table (4) it can draw all output of neural network versus number of tested voice command to show the recognized command among five voice command. From table (4) and Figure (8), output "Y1" represent the voice command "GO".

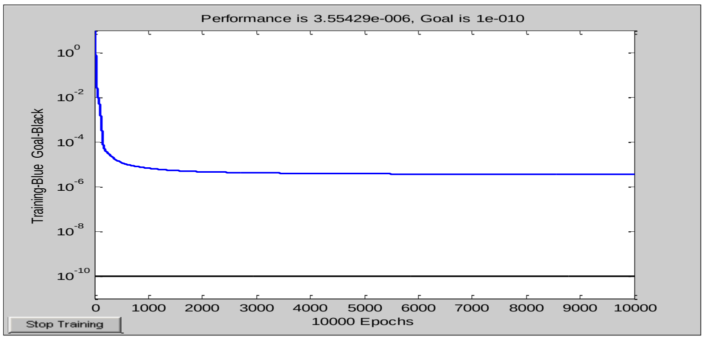

Figure (7), Training Process Five Control Output

when the value of "Y1" is number (1) which represents the output (1) is active and the number (-1) represents the output is disable. The output of the test may not be 1 or -1 but may reach these values. Therefore the negative value (-0.9996) give wrong voice command, from "Y3" it is noted that it's value positive is (0.3218) that mean the wrong voice command is "BACK". This procedure is done for all command that is shown in figures below:

Input speech	Output					Recognized command
	Y1	Y2	Y3	Y4	Y5	
"Go1"	-0.9996	-1.0000	0.3218	-0.9939	-0.8297	"Back"
"Go2"	0.9984	-0.9982	-0.9999	-0.9989	-0.9988	"Go"
⋮	⋮	⋮	⋮	⋮	⋮	⋮
"Go10"	0.9999	-0.9992	-1.0000	-0.9973	-0.9981	"Go"
"Left1"	-0.9952	0.9944	-1.0000	-0.9968	-1.0000	"Left"
"Left2"	-0.9990	0.9973	-1.0000	-1.0000	-0.9985	"Left"
⋮	⋮	⋮	⋮	⋮	⋮	⋮
"Left10"	-1.0000	0.9942	-1.0000	-1.0000	-0.9994	"Left"
"Back1"	-0.9999	-1.0000	0.9993	-0.9989	-0.9959	"Back"
"Back2"	-1.0000	-1.0000	0.9958	-0.9998	-0.9974	"Back"
⋮	⋮	⋮	⋮	⋮	⋮	⋮
"Back10"	-1.0000	-1.0000	0.9990	-0.9995	-0.9982	"Back"
"right1"	-0.9999	-1.0000	-0.9988	0.9989	-0.9998	"right"
"right2"	-0.9967	-0.9999	-0.9998	0.9951	-0.9952	"right"
⋮	⋮	⋮	⋮	⋮	⋮	⋮
"right10"	-0.9990	-0.9983	-0.9991	0.9972	-1.0000	"right"
"stop1"	-0.9910	-0.9993	-0.9943	-0.9887	0.9978	"stop"
"stop2"	-1.0000	-1.0000	-0.4854	-0.9742	0.9907	"stop"
⋮	⋮	⋮	⋮	⋮	⋮	⋮
"stop10"	-0.9965	-0.9995	-0.9959	-0.9997	0.9928	"stop"

Table (4), The Result of Neural Network Training (continued)

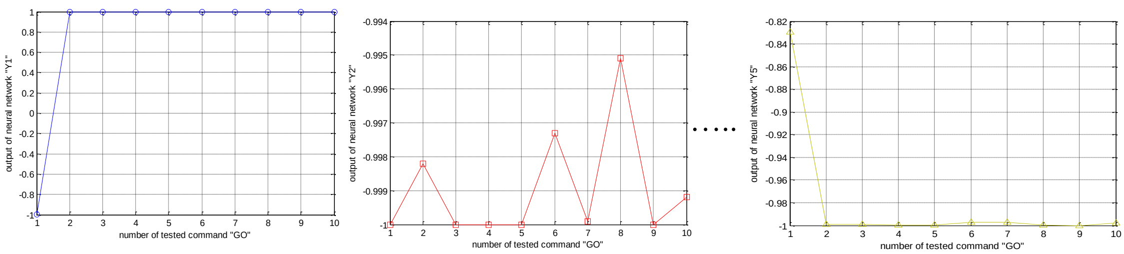

Figure (8), Output of Neural Network for Command "Go"

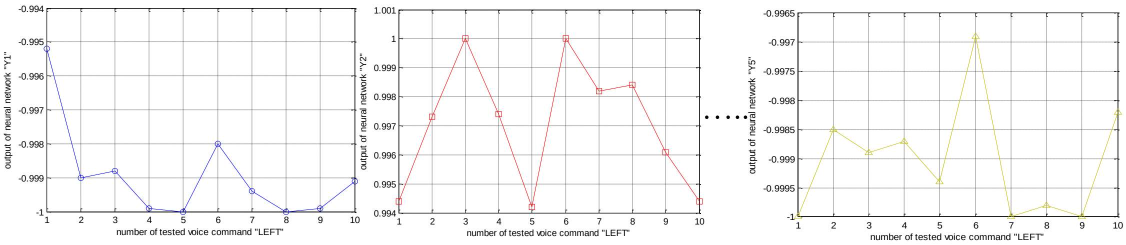

Figure (9), Output of Neural Network for Command "LEFT"

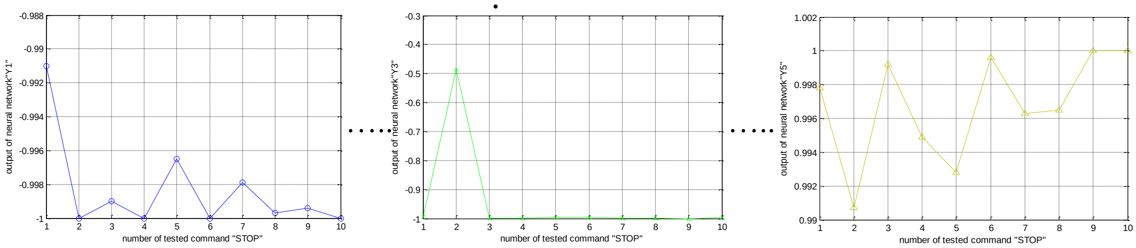

Figure (10), Output of Neural Network for Command "STOP"

6. Experimental Work

The wheelchair that used in this work has three connecting rod (one in front and two in rear of wheelchair) that connect the two sides of wheelchair; each rod has joint in middle this will enable the wheelchair to be portable. The wheelchair is 65 cm (25.5 inches) wide and 127 cm (50 inches) long, measured with the rear caster extended fully behind the chair. The front of the tray is 76 cm (30 inches) from the ground, and the base of the seat is 51 cm (20 inches) above the ground. The robotic wheelchair used in this work shown in Figure (2) was built by the BEG company (British company) with joystick.

Figure (2), Rear projection of used wheelchair

6-1 Weight of Wheelchair

The weight of wheelchair alone is 17kg, while weight of wheelchair with two motors and with experimental circuit was 20kg. Weight of person that sits down inside the wheelchair and which was conducted upon him practical experiences was 65kg. The weight of wheelchair with the person sit down inside the wheelchair was 85kg. Maximum weight that the wheelchair can supported with respect to the person that sit down inside it and can give the same work and performance is approximately 80kg, and then the maximum weight for the wheelchair with the person approximately 100kg.

I make experimental work on wheelchair that has the person sit down inside it has weight was 65kg and this wheelchair shown in figure (2) and this practical experience gives excellent result and this is illustrated in table (5). From this table I conclusion that the proposed method better than the previous works.

6-2 Wheelchair Battery

The battery used in this work is wet type .Wet batteries use the chemical reaction between lead and sulphuric acid to create electrical energy. As the batteries need filling with distilled water, they do have a higher maintenance rate, but are lighter than Gel or AGM (Absorbed Glass Mat) batteries.

6-3 Wheels

Wheelchair has four wheels, two rear wheels and two castor wheels, the two caster wheel are fixated in wheelchair base in front all wheels have the same diameter (18 cm) .The drive wheels are in rear on either side of the base, allowing the chair to turn according to voice command, wheels engages directly to a gear train that transmit torque form motor to wheels by two grooves in each wheel and nut.

6-4 Motors

Motors are arguably one of the most important parts of a mobile robotics platform. Overpowered motors cause inefficiency and waste the already limited supply of power from the on-board batteries, while undersized motors could be short on torque at critical times. The optimal rotation speed and the available speed range of the motor must also be taken into consideration. Too high of an output rpm from the motor shaft will cause the robot to operate at a fast, uncontrollable speed. Too low of an output and the robot will not be able to attain a suitable speed to meet the user's needs. The torque output of the motor also plays a role in the performance because if the torque is not sufficient, locomotion may not occur in certain situations. Therefore, much consideration was put into the selection of the proper motor for the platform [Philips, 2003].

Motors come in many shapes and sizes. There are electromagnetic direct current (DC) motors and electromagnetic alternating current (AC) motors and a number of variations of each. AC motors are typically used for large applications, such as machine tools, washers, dryers, etc., and are powered by an AC power line. Since the typical power supply for mobile robotic is a DC battery, and technology for transforming DC to AC is very expensive in both terms of monetary cost and power cost, AC motors where ruled out as an option for the robot.

DC motors are commonly used for small jobs and suited the purposes of the platform very well. Figure (3) shows the 12V DC motor use in wheelchair.

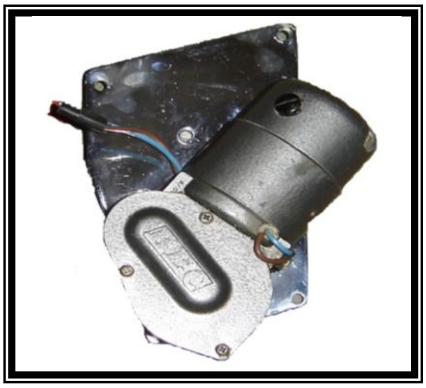

Figure (3), DC motor used in wheelchair

7. Hardware Components Added to Original Wheelchair

The modification that adds to original wheelchair with removing joystick that designs before (to modifying wheelchair function according to person injury especially for the injuries who suffer from spasms and paralysis of extremities) makes its physical design very effective. It is a combination of various physical (hardware) and computational (software) elements that integrate the subsystems of the wheelchair to work in one unit. In terms of hardware components the main components that added to wheelchair are interfacing circuit, microphone (headset microphone) and notebook computer (host computer).

7-1 Microphone

A quality microphone is the key when utilizing automatic speech recognition (ASR). In most cases, a desktop microphone just will not do the job. They tend to pick up more ambient noise that gives ASR programs a hard time. Hand held microphones are also not the best choice as they can be cumbersome to pick up all the time. While they do limit the amount of ambient noise, they are most useful in applications that require changing speakers often, or when speaking to the recognizer is not done frequently (when wearing a headset is not an option).

The best choice and by far the most common is the headset style. It allows the ambient noise to be minimized, while allowing you to have the microphone at the tip of your tongue all the time [Cook, 2002]. Headsets are available without earphones and with earphones (mono or stereo) in this work the headphone type (FANCONG FC-340) is employed.

7-2 Relay Driver Interfacing Circuit

A relay can be used to switch higher power devices such as motors and solenoids. If desired, the relay can be powered by a separate power supply, so, for instance, 12V motors can be controlled by the parallel port of notebook computer. Free welling diode can be used to protect the relay contact and prevent damage to the transistor when the relay switches off.

An intermediate stage between control signal (output of parallel port) and motors consists of a combination of component relays, transistors, diodes, capacitors, resistors and buffer 74ABT245 as shown in Figures (4) and (5), it uses to protect parallel port against any expected damage, The 74ABT245 high-performance BiCMOS device combines low static and dynamic power dissipation with high speed and high output drive shown in Figure (6).

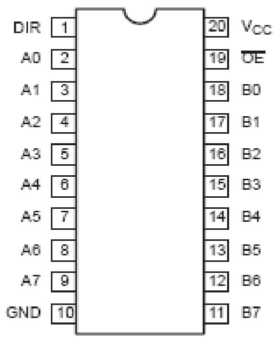

Figure (6), Pin Configuration of 74ABT245

Figure (4), Relay Interfacing Fixated on Wheelchair

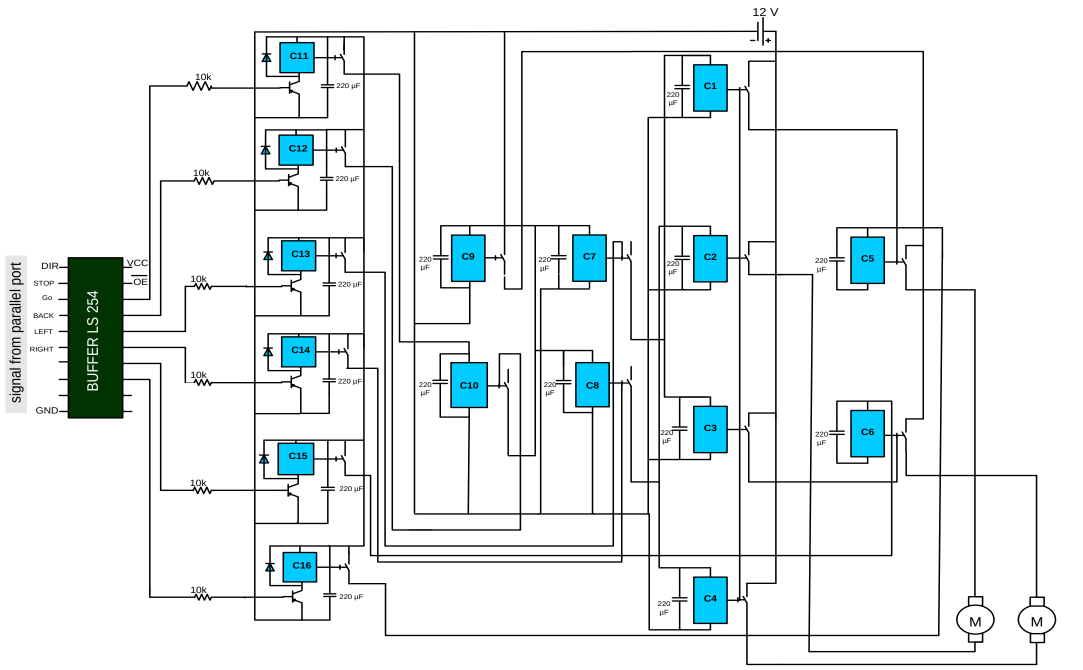

Figure (5), Relay Interfacing Circuit

The 74ABT245 device is an octal transceiver featuring non-inverting 3-State bus compatible outputs in both send and receives directions. The control function implementation minimizes external timing requirements. The device features an Output Enable (OE) input for easy cascading and a Direction (DIR) input for direction control [buffer 63]. Descriptions and function of the 74ABT245 are shown in tables (1) and (2).

PIN NUMBER	SYMBOL	NAME AND FUNCTION
1	DIR	Direction control input
2, 3, 4, 5, 6, 7, 8, 9	A0 – A7	Data inputs/outputs (A side)
18 ,17, 16, 15, 14, 13, 12, 11	B0 – B7	Data inputs/outputs (B side)
19	$\overline{\text{OE}}$	Output enable input (active-LOW)
10	GND	Ground (0 V)
20	VCC	Positive supply voltage

Table (1), Pin Description of 74ABT245

$\overline{\text{OE}}$	DIR	An	Bn
L	L	An = Bn	Inputs
L	H	Inputs	Bn = An
H	X	Z	Z

Table (2), Function Table of 74ABT245

H = High voltage level
L = Low voltage level
X = don't care
Z = High impedance “off” state

7-3 Useful Matlab Language Tool in Experimental Work

Now it must show how voice can control the wheelchair, it begins at microphone using matlab program language.

y = wavrecord(n,Fs) (12)

record n samples of an audio signal, sampled at a rate of Fs Hz (samples per second). The default value for Fs is 11025 Hz. Standard sampling rates for PC-based audio hardware are 8000, 11025, 22050, and 44100 samples per second. Stereo signals are returned as two-column matrices. The first column of a stereo audio matrix corresponds to the left input channel, while the second column corresponds to the right input channel [Matlab Help, Version 7]. In this work mono type 22050 sample per second presented that means single column of audio matrix.

This single column represents isolated word command like GO, LEFT ……. etc will pass through Multiridgelet transform to compare the uttered word (command) with data base (saved as neural network weights), when the comparison processes are successful, now use this command to control wheelchair by parallel port of host computer using matlalb program language.

DIO = digitalio('adaptor',ID) (13)

DIO = digitalio('adaptor',ID) creates the digital I/O object DIO for the specified adaptor and for the hardware device with device identifier ID. ID can be specified as an integer or a string [Matlab Help, Version 7].

'adapter' The hardware driver adaptor name. The supported adaptors are advantech, keithley, mcc, nidaq, and parallel.

ID The hardware device identifier.

DIO The digital I/O object.

lines = addline(obj,hwline,'direction') (14)

Lines = addline(obj,hwline,'direction') adds the hardware lines specified by hwline to the digital I/O object obj. direction configures the lines for either input or output. Lines are a row vector of lines [Matlab Help, Version 7].

obj A digital I/O object.

hwline The numeric IDs of the hardware lines added to the device object. Any MATLAB vector syntax can be used.

'direction' The line directions can be In or Out, and can be specified as a single value or a cell array of values.

Finally the recognized voice commands shall convert to binary vector using matlab program language putvalue(obj,data) as shown below in table (3).

putvalue(obj,data) writes data to the hardware lines contained by the digital I/O object obj [Matlab Help, Version 7].

obj A digital I/O object.

data A decimal value or binary vector.

Recognized Voice Command	Binary Vector Input /74ABT245				
GO	1	0	0	0	0
LEFT	0	1	0	0	1
RIGHT	0	0	1	0	1
BACK	0	0	0	1	0
STOP	0	0	0	0	1

Table (3), Voice Command Corresponding Binary Vector

8. Results Simulation

8-1 Results of the Proposed Algorithm

For the proposed algorithm, 50 utterances for each speech command are used for training step. Table (5) shows the result of applying 2-D discrete Multiridgelet transformfor for different data. These data consist of five voice command and each command has 50 utterances. In table (5), the first column represents the training (reference) data with specific sequence of commands that are namely (command1-command5). For example, the first row in table (5) consists of 50 utterances for command "GO". The second column represents the testing data for the same command but for different speaker. In the same row (i.e. first), the test data have 10 utterances, as a test for command "GO".

The next column represents the recognition rate for applying 2-D discrete Multiridgelet transform for different data (i.e. represents the speech command recognition). So in the first row (9 /10) shows 9 from 10 tested commands that can be recognized correctly. Thus, the rate of recognition command "GO" is 90%, while by using Multiwavelet transform is 70%. The last two columns represents the rate for all commands in percentage 100% by using Multiridgelet and Multiwavelet transform and the overall percentage rate was 98%, 82% for two transforms respectively. From this table I conclude that the proposed method is better than the other method.

No.	Reference sequence (command1-command5)	Testing sequence	Multiridgelet Transform	Multiwavelet Transform	Percentage in 100% for Multiridgelet Transform	Percentage in 100% for Multiwavelet Transform
1	go1-go50	go1-go10	9/10	7/10	90%	70%
2	left1-left50	left1-left10	10/10	8/10	100%	80%
3	back1-back50	back1-back10	10/10	10/10	100%	100%
4	right1-right50	right1-right10	10/10	8/10	100%	80%
5	stop1-stop50	stop1-stop10	10/10	8/10	100%	80%

Tables (5), The Results of Multiridgelet, and Multiwavelet Transform

8-2 Experimental Results

To examine the performance of the proposed algorithm, some experimental tests were done by applying different type of voice command experimentally like (go, back, right …etc). The following observation has been obtained from it where; many cases for the wheelchair motions are considered, showing the path that the wheelchair would take on its motion, making use of different voice command to steer wheelchair in proper direction.

8-2-1 Linear Path

Linear path of wheelchair can be obtained by single isolated voice command "GO" and "BACK" according to specified direction that user recommended, Figure (11) shows linear motion of wheelchair, motion of wheelchair either in positive or negative direction of x-axis when the rear (driving) wheels of wheelchair have the same direction of rotation (both wheels are rotate in clockwise or anti-clockwise) and have the same magnitude (V1=V2).To stop motion of wheelchair "STOP" voice command is used.

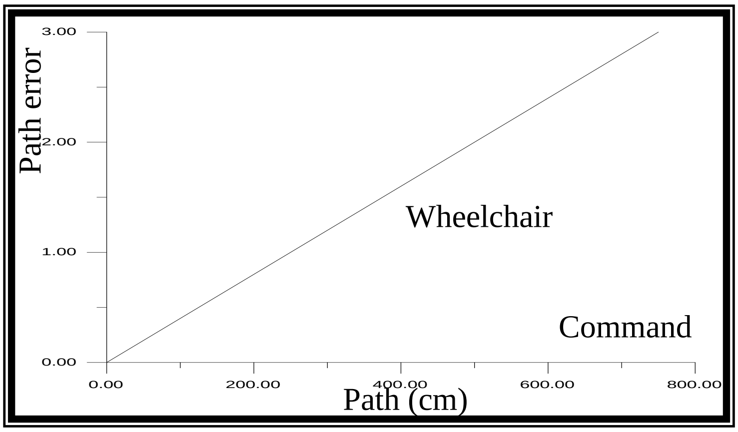

Figure (12), Actual Wheelchair Path in Linear Motion

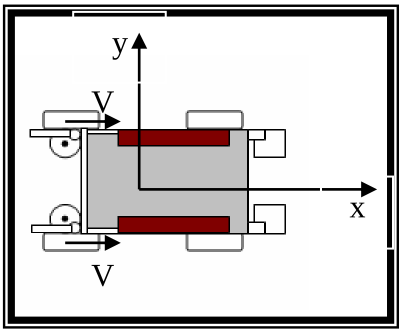

Figure (11), Linear Path Corresponding to "GO" Voice Command

Figure (12) shows error path of wheelchair from the commanded path which represents in figure the x-axis.

8-2-2 Circular Path

Circular path of wheelchair can be obtained by single isolated voice command either "LEFT" or "RIGHT" according to specified direction that user recommended, Figure (13) shows rotational motion of wheelchair in clockwise using voice command "RIGHT" (V1=0, V2= not zero).

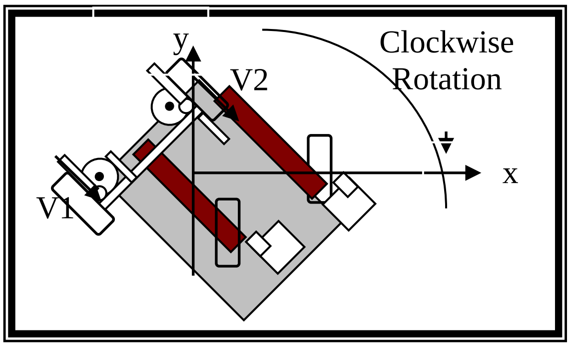

Figure (13), Rotation Corresponding to "RIGHT" Voice Command

These voice commands generate circular path, To stop rotation motion of wheelchair "STOP" voice command is used.

9. Conclusion

In this paper a method of isolated word speech recognition system was proposed to control the wheelchair and therefore make proposed technique be more efficient in real time operation used in control of mobile robot.

This paper present a proposed 2-D Multiredgelet transform computation method that verifies the potential benefit of Multiwavelet and gain a much improvement in term of low computation complexity.A single level decomposing in the Multiwavelet domain is equivalent to scalar wavelet decomposition. Thus, although computing complexity is double for DMWT compared to DWT.

Discrete Multiwavelet transform computation algorithm using repeated row preprocessing should be applied to matrix with a size at least to 4×4. Multiredgelet transform is important technique in word recognition application due to eliminate the noise, sharpening or smoothing the speech. A nonlinear enhancement function was applied on the discrete Multiredgelet coefficients; it has good result (enhancement speech) over a linear enhancement function in multiple speeches. For Multiredgelet gaining algorithm (by applying linear or nonlinear enhancement functions), the success of Multiredgelet thresholding-feature extraction method is because of:

a. Multiresolution: Multiwavelet transform analyzes a signal at a nested set of scales.
b. Locality: Multiwavelet transform coefficients are localized simulta-neously in frequency and angles; hence Multiredgelet can match a wide range of signal components, from transients to harmonics.
c. Ease of implementation: Multiredgelet transform provides a more direct way-simply of use and ease of implementation, since the geometrical structure of the best sequence of directions is the same in forward and inverse best sequence of directions.

This rate of recognition is sufficient to control the wheelchair in safe without using additional sensor like (ultrasonic, infrared ….etc); all these sensors are used to support the wheelchair when the recognition rate is weak.

10. References

Cook S., 2002,"Speech Recognition How To", Revision v2.0 April 19, 2002.

Hosseini E., Amini J., Saradjian M.R., 1996, "Back Propagation Neural Network for Classification of IRS-1D Satellite lite Images", Tehran University, 1996.

Hrnčár M., 2007, " Voice Command Control For Mobile Robots", Department of Control and Information Systems Faculty of Electrical Engineering, University of Žilina, Žilina 2007.

Komiya K., Morita K. , Kagekawa K. and Kursosu K., 2000, "Guidance of a Wheelchair by Voice " *proceeding of IEEE*,PP102-107,2000.

Li J., Pan Q., Zhang H. and Cui P., 2003, "Image recognition using Radon Transform", *Proceeding of IEEE*, 2003.

Lim S. and Jeon J., 1997, " Multiple Mobile Robot Interface Using a Low Cost Voice Recognition Chip ", Department of Mechatronics, School of Electrical and Computer Engineering, *proceeding of IEEE*, PP 402-407,1997.

Mahoney R., 1997, " Robotic products for rehabilitation: Status and strategy", *Proceeding of IEEE,* 1997.

MATLAB Help, Version 7.

Minh N. and Vetterli M., 2003, "The Finite Ridgelet Transform for Image Representation", *Proceeding of IEEE,* January, 2003.

Moon I., Lee M., Ryu J. and Mun M, 2003. "Intelligent Robotic Wheelchair with EMG, Gesture, and Voice-based Interfaces", *Proceeding of IEEE,* PP 3453-3458,October 2003.

Philips semiconductor, 2003, "Integrated circuit data sheet", Feb 06.

Ram Kumar P., Ramana Murthy M.V., Eashwar D. and Venkatdas M., 2005, " Time Series Modeling Artificial Neural networks", *Journal of Theoretical and Applied Information Technology*, 2005.

Rockland R. and Reisman S., 1998, "Voice Activated Wheelchair Controller", in proceeding of IEEE, PP 128-129,1998.

Sajkowski M., 2002, "Voice Control of Dual-Drive Mobile Robots-Survey of Algorithms", Silesian University of Technology, Institute of Theoretical an Industrial Electronics, *Proceeding of IEEE,* November 9-11, 2002.

Simpson R. and Levine S., 2002, "Voice Control of a Powered Wheelchair", IEEE Transaction on Neural System and Rehabilitation Engineering , Vol. 10, No. 2, PP 122-125,June 2002.

Sundeep, Portia, Singh H., Mobin A. and Agrawal S., 2000," Design and Development of Voice-Cum-Auto Steered Robotic Wheelchair Incorporating Reactive Fuzzy Scheme for Anti-Collision and Auto Routing", in proceeding of IEEE, PP 192-195, 2000.

Takiguchi T., Yamagata T., Sako A., Miyake N., Revaud J. and Ariki Y., 2008, " Human-Robot Interface Using System Request Utterance Detection Based on Acoustic Features", International Journal of Hybrid Information Technology Vol. 1, No. 3, July, 2008.

Terrades O. and Valveny E., 2003," Radon Transform for Lineal Symbol Representation", *Proceeding of IEEE of the Seventh International Conference on Document Analysis and Recognition (ICDAR'03),* 2003.

Valin J., Yamamoto S., Rouat J., Michaud F., Nakadai, K. and Okuno H., 2007, "Robust Recognition of Simultaneous Speech by a Mobile Robot", *Proceeding of IEEE* ,VOL. 23, NO. 4, AUGUST 2007.

Yanco H., 2000, "Shared user-Control of a Robotic Wheelchair System", *Ph.D. Thesis*, *Massachusetts Institute of Technology*, 2000.

Zurada J., 1996, Introduction *to Artificial Neural Network,* Jaico Publishing house, 1996.