

Silla de Ruedas Controlada por Voz

Autores: Montoto Paredes, Jesús Fernando y Gutiérrez Garrido, Antonio

Universidad Politécnica de Pachuca
Ingeniería Mecatrónica
Carretera Pachuca-Cd. Sahagún, km 20, Ex-Hacienda de Santa Bárbara, Municipio de Zempoala, Hidalgo,
Tel.01 743 791 18 00 Extensión 2213
Correo electrónico. fer_19335@hotmail.com

Resumen

En el presente proyecto muestra una silla de ruedas con control autónomo guiado por voz, incluyendo el sistema robótico de movimiento y el sistema de procesamiento y reconocimiento de voz, los que conjuntan un sistema totalmente mecatrónico. El usuario emitirá el comando de voz para que la silla realice el movimiento direccional a donde se desea dirigir; los comandos de voz básicos son: “adelante”, “atrás”, “izquierda”, “derecha”, “alto”. La ventaja de esta silla de ruedas es que contribuirá a mejorar la calidad de vida y las posibilidades de movimiento autónomo de las personas con capacidades diferentes.

Introducción

Este proyecto tiene como objetivo desarrollar una silla de ruedas automática guiada a través de comandos de voz, para que esta sirva de asistencia a personas cuadraplégicas o con alguna capacidad diferente, para permitirles un mayor grado de autonomía en lo que a su movilidad se refiere. Y es así como este proyecto tiene como finalidad mejorar substancialmente la calidad de vida de las personas con limitaciones físicas para realizar desplazamientos sin apoyo humano, al

desarrollar esta nueva tecnología basada en comandos de voz como interfase hombre-maquina e innovación tecnológica en el ámbito de la ingeniería aplicada y mecatronica. Crear un software capaz de reconocer la voz.

Este prototipo tiene como finalidad principal elaborar una silla de ruedas con control autónomo guiado por voz, incluyendo el sistema automático de movimiento y el sistema de procesamiento y reconocimiento de voz, los que conjuntan un sistema totalmente mecatrónico. El usuario entrenara los comandos y emitirá el comando de voz deseado para que la silla realice el movimiento direccional a donde se desea desplazar; los comandos de voz básicos son: “adelante”, “atrás”, “izquierda”, “derecha”, “alto”. La ventaja de esta silla de ruedas es que contribuirá a mejorar la calidad de vida y las posibilidades de movimiento autónomo de las personas con capacidades diferentes.

Con el avance de la tecnología el hombre ha buscado la manera de comunicarse con las maquinas, desde otorgarles “inteligencia”, capacidad de procesar información, hasta proporcionarles visión periférica, o una comunicación con el entorno físico, en la actualidad se han

desarrollado interfaces hombre-maquina aplicadas a sistemas de seguridad, ergonomía y procesos de manufactura, que han impulsado el desarrollo de nuevas tecnologías tales como, software de reconocimiento de voz.

Para poder entender la estructura con la cual decidimos echar a andar este prototipo se decidió dividir el proyecto en las siguientes etapas:

Software de reconocimiento

Etapas mecánica

Electrónica de potencia

Etapas de control

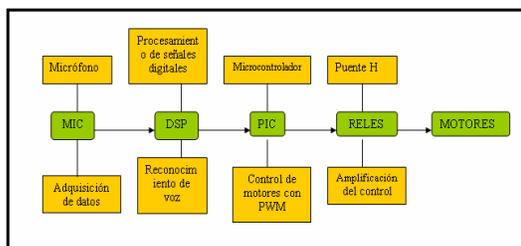


Fig. 1 Diagrama de bloques del sistema.

Marco Teórico

Justificación

En la actualidad existen infinidad de afecciones que impiden el desplazamiento autónomo de las personas, suelen depender de aparatos para usos especiales como son las sillas de ruedas, en algunos casos severos de discapacidad estos

aparatos no son suficientes o adecuados, un buen ejemplo es la cuadriplejía en la que los enfermos carecen totalmente de sus facultades motoras entonces las sillas convencionales tienen la ventaja de funcionalidad automatizada con la ayuda de un joystick e incluso en la mayoría de los casos solo es mecánica, pero en nuestro proyecto se está dando un valor agregado a estos aparatos controlándolos por emisión de voz.

En recientes investigaciones hemos recopilado información suficiente para aventurarnos a decir que actualmente en el mercado no se encuentra a la venta una silla de ruedas o algún tipo de aparato con las cualidades que presenta nuestro prototipo, este muestra una importante innovación en el medio de la ingeniería y en el campo de la electrónica biomédica.

Existen un sin número de afecciones las cuales deterioran el movimiento motriz del cuerpo, parcialmente hasta totalmente, un claro ejemplo es la cuadriplejía, en este tipo de discapacidad el paciente pierde el movimiento de sus cuatro partes motrices, adquiriendo un grado de discapacidad severo en donde sillas de ruedas convencionales no resuelven la problemática de su desplazamiento autónomo, en algunos casos se optaba por implementar una silla de ruedas automatizada, pero guiada por un joystick, siendo esta una impotencia para el usuario.

Cuando un paciente es afectado por algún tipo de padecimiento o discapacidad que le impida total o parcialmente llevar a cabo satisfactoriamente el desplazarse con total autonomía motriz, se tendrá una solución con la puesta en marcha de este prototipo, siendo así una excelente opción para personas con este tipo de padecimientos.

Objetivos

En la culminación de este proyecto tenemos esperado la creación de una silla de ruedas capaz de transportar de manera autónoma y como principal medio la voz a personas con capacidades diferentes.

Desarrollar nueva tecnología como interfase hombre-maquina en el que la voz juega un papel preponderante.

Que este proyecto sirva como herramienta de tesis para sus autores.

Posicionar en el mercado el proyecto para su comercialización.

Obtener una patente.

Desarrollo del proyecto

Software de reconocimiento

Generalmente para analizar fenómenos físicos, uno se basa en las señales que ellos producen. El dominio natural de la señal es el tiempo, o el espacio si se trata de una imagen. Sin embargo, para visualizar atributos particulares o para realizar procesos de una manera más simple, desde el siglo pasado se ha intentado transformar la señal y llevarla a otros dominios como por ejemplo la frecuencia. El primero de estos intentos corresponde al trabajo de Fourier.

(Trina Adrián de Pérez, 1994)

El análisis tradicional de Fourier no es útil cuando la señal a analizar carece de una regularidad estadística determinada o cuando tiene variaciones locales de la frecuencia. Se piensa entonces que si se

escogen porciones de la señal o lo que es lo mismo, se hacen pasar los datos por una ventana previa al cálculo de la transformada de Fourier, es posible obtener la localización en tiempo de cada componente frecuencial. Esto representa observar la señal $x(t)$ a través de la ventana $w(t)$ trasladada hasta el tiempo t , que es lo que hace la Transformada de Fourier de Corto Plazo o STFT (Short Time Fourier Transform) de la señal $x(t)$, es decir (1):

$$STFT_X(t, f) = \int_{-\infty}^{\infty} x(\tau) w^*(\tau - t) e^{-j\omega\tau} d\tau$$

(1)

$$STFT_X(t, f) = e^{-j\omega t} \int_{-\infty}^{\infty} X(f') W^*(f' - f) e^{j2\omega t f'} df'$$

(Kadambe & Boudreaux-Bartels. 1992).

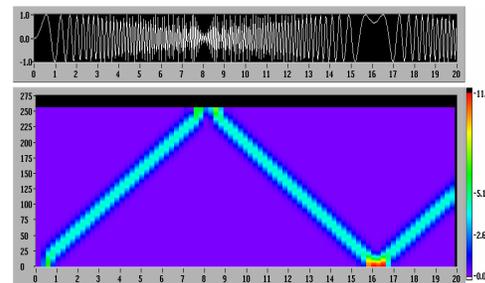


Figura 2. Función de la SFTF
 (Transformada de Fourier de Corto Plazo)

En parte superior de la figura 2 se muestra una señal de frecuencia variable, y magnitud constante, en la parte de abajo se muestra la Transformada de Fourier de Corto Plazo, en la cual se puede observar como la señal incrementa su frecuencia en Y a medida que el tiempo en X avanza,

hasta un punto máximo de 250Hz y posteriormente decrece a cero en Y, la magnitud de la señal se puede observar en el color del espectrograma.

Se puede ver a la STFT como la proyección de la señal $x(t)$ sobre una función base sinusoidal limitada temporalmente a través de una ventana fija $w(t)$ y, debido a que la operación es invertible, esto al mismo tiempo equivale a representar la señal $x(t)$ a través de una familia de funciones bases del tipo: $w^*(\tau - t)e^{-j\omega\tau}$. Otra forma de verlo es como un filtraje de la señal $x(t)$ con un filtro pasabanda de ancho de banda constante trasladado a las diferentes frecuencias de análisis

Una vez elegida la ventana $w(t)$, la resolución para todo el plano tiempo-frecuencia $t-f$ estará fija. Por lo tanto para una STFT dada podemos aumentar la resolución temporal con funciones más estrechas en tiempo con lo que estaremos perdiendo resolución en frecuencia ya que los filtros se hacen más anchos. El mejor intento fue el de Gabor quien propuso el uso de una ventana con envolvente gausseana lográndose una buena resolución en el dominio $t-f$. Pero aún así se sigue dependiendo del tamaño de la ventana elegida.

La característica principal de este enfoque es que usa un marco Matemático bien definido y que establece representaciones consistentes de los patrones de voz que pueden usarse para comparaciones confiables a partir de un conjunto de muestras rotuladas, usando algoritmos de entrenamiento. La representación de los patrones de voz puede ser una plantilla

(template), o un modelo estadístico (HMM: Hidden Markov Model), que puede aplicarse a un sonido (más pequeño que una palabra), una palabra, o una frase.

La técnica de reconocimiento consiste básicamente en dos pasos:

Primer Paso: entrenamiento de patrones

Segundo Paso: comparación de patrones

Entrenamiento de patrones

En esta etapa se construye un patrón de referencia asociado a cada palabra (o sub-unidad de palabra) que se quiere reconocer, basándose en los vectores característicos de todas las palabras usadas para el entrenamiento. Hay varias formas en que puede realizarse el entrenamiento.

1. Entrenamiento casual: un único patrón de sonido es usado para crear el patrón de referencia correspondiente o un modelo estadístico Aproximado.

2. Entrenamiento robusto: se utilizan varias versiones de cada palabra o sonido a reconocer (generalmente provenientes de un sólo locutor) para construir un patrón de referencia promedio o un modelo estadístico promedio.

3. Entrenamiento por clustering: se utiliza un gran número de versiones de cada palabra o sonido (provenientes de un gran número de locutores) para construir patrones de referencia o modelos estadísticos más confiables.

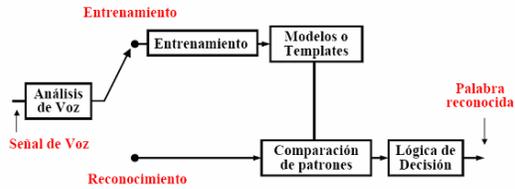


Fig. 3 Reconocimiento de voz basado en comparación de patrones.

COMPARACION DE PATRONES

En la etapa de comparación de patrones se realiza una comparación directa entre el vector característico asociado a la señal de voz desconocida (a reconocer) y todos los posibles patrones aprendidos en la etapa de entrenamiento, de manera de determinar el mejor ajuste de acuerdo a algún criterio. Surge la necesidad de definir una medida de similitud (distancia) entre vectores característicos que permita determinar cuál es el patrón de referencia que mejor se ajusta a la señal a reconocer.

En general, debido a que una misma palabra es emitida con diferentes velocidades cada vez que es pronunciada, para poder realizar la comparación es necesaria una normalización temporal, previa a la extracción de característica. El enfoque más simple sería una transformación lineal del eje tiempo, pero esto no es realista y en la práctica se recurre a técnicas de programación dinámica (Dynamic Time Warping).

Primera versión del programa de reconocimiento de voz implementado

La figura 2 muestra el programa de reconocimiento de voz implementado en la silla de ruedas, este programa tiene la capacidad de ser entrenado para

reconocer diferentes mandos (palabras). En especial tiene en el punto VI la capacidad de reconocer tres diferentes tipos de “Listas” de palabras.

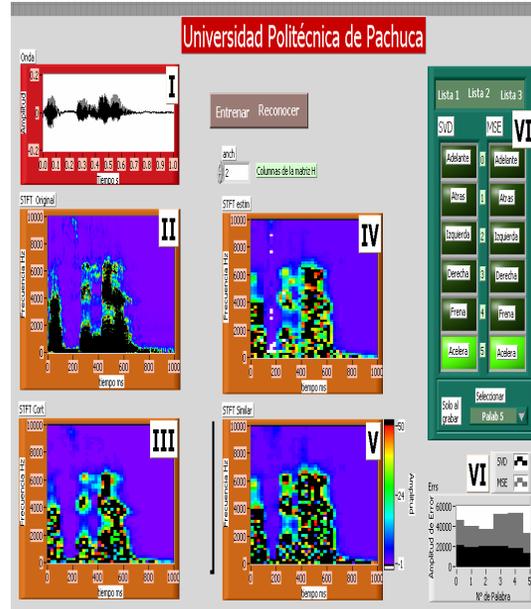


Figura 4 Programa de reconocimiento de voz.

- I Adquisición del sonido por medio del micrófono de la computadora personal a una tasa de muestreo de 8000 muestras/segundos.
- II Procesa la señal digital con la “Transformada de Fourier de Corto Plazo” (STFT Short Time Fourier Transform), la cual es una matriz que muestra en X el tiempo, en Y el espectro de frecuencias que contiene el lapso de la señal y en Z por intensidad de colores, la magnitud de las frecuencias.
- III Se hace una reducción de dimensiones en X (de 201 a 40) y Y (de 501 a 50) del espectrograma, así como reducción de los valores de intensidades en Z (de 65536 a 256

colores), dando como resultado una matriz 201,501, 65536, a 40, 50, 256 (X,Y,Z).

- IV** Se estima el espectrograma más aproximado a III, basándose en la base de datos almacenada previamente. Para lograr esta estimación se utiliza el método “Descomposición del Valor Singular” (SVD Singular Value Decomposition).
- V** Con base al espectrograma estimado en IV, se busca en la base de datos el más parecido.
- VI** Se muestran dos métodos utilizados para la identificación de voz SVD y MSE (Mean Square Error), teniendo más precisión el método SVD que MSE.
- V** Con base al espectrograma estimado en IV, se busca en la base de datos el más parecido.

Etapa mecánica

Sistema de transmisión

Debido a que contamos con una silla de ruedas manual, se elaboró un diseño, en el que se montaron 2 motores de CD, colocando un sistema de transmisión a través de un juego de engranajes acoplados con una flecha, un motor para cada rueda, la flecha es la encargada de transmitir movimiento a través de los engranes cónicos a las ruedas traseras para generar el movimiento, para llegar a esto se elaboro el siguiente diseño mecánico:



Fig. 5 Diseño de soporte para los motores.

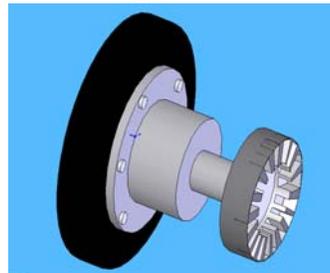


Fig. 6 Diseño de transmisión.



Fig. 7 Diseño final del sistema mecánico.

Existen criterios para limitar el peso máximo en una silla de ruedas eléctrica, dependiendo del material con el que se fabrica la estructura de la silla, ensamblajes uniones etc., la masa máxima para el usuario que se considero fue de 108 Kg. Por ser especificado en la silla de ruedas utilizada.

$$PT = PEM + PMU + PFP + PSC$$

PEM Peso de estructura mecánica = 490.0N

PMU Peso máximo del usuario = 1058.0N

PFP Peso máximo de la fuente de poder = 390N

PSC Peso del sistema de control = 19.6N

PT Peso total = 1957.6N

Sobre requerimientos de diseño en espacios para sillas de ruedas se norman según los siguientes aspectos:

TMSD Tramo máximo sin descanso = 9m

PMI Pendiente máxima en interiores = 11%, Angulo máximo 6°

PME Pendiente máxima en exteriores = 8%, ángulo máximo de 4.6°

PMR Pendiente máxima recomendable = 6%, ángulo máximo 3.4°

Se analiza la fuerza en una pendiente con la mayor inclinación ascendente, se toma un coeficiente de fricción sintético de $0.57 \mu c$ en materiales de hule sobre concreto, y con el peso máximo a movilizarse se calcula la fuerza máxima necesaria para movilizarse hacia delante, atrás, derecha o izquierda ya que la fuerza requerida F para el movimiento debe ser igual a la fuerza de rozamiento Fr , se toma la normal en función del Angulo β que representa el 11% del PMI

$$N = m.g.\cos \beta$$

La fuerza máxima F necesaria para la movilización es la suma de la fuerza de rozamiento Fr con la componente de peso ($m.g.\sen \beta$)

$$Fr = \mu N$$

$$F = Fr + m.g.\sen \beta$$

$$F = (m.g)(\mu \cos \beta + \sen \beta)$$

SISTEMA DE ALIMENTACION

Para la selección correcta de la fuente de poder que alimentara nuestro prototipo, se llevaron acabo los siguientes cálculos:

Características de las baterías:

Voltaje= 12V

Corriente= 5Ah

1 Hp=746 W

¼ Hp=186.5 W

(186.5W) (2motores) = 376 W

(12V) (11baterias) = 132V

Funcionamiento optimo del sistema

$$\frac{376W}{132V} = 2.84A$$

La alimentación mínima en el sistema debe de ser de 8 baterías, debido a lo siguiente:

$$376W/96V=3.9A$$

Los dispositivos electrónicos de potencia que aplicamos en este prototipo no soporta mas de 5 A.

DURACION DE ENERGIA

El tiempo estimado con 11 baterías (2.84A) fue de 1 hora y 45 minutos, se calculo como se muestra a continuación:

$$\frac{(60 \text{ min})(5A)}{(2.84A)} = 1.76 = 1.45 \text{ Hr.}$$

ELECTRONICA

Para el control de los motores se implemento la tarjeta que se muestra en la figura 8. Consta de dos puentes H, uno para cada motor, cada puente esta elaborado con cuatro relevadores, estos dispositivos no soportan la corriente de la fuente de poder (132V), por lo tanto se pondrán relevadores de capacidad adecuada comandados por los relevadores de la tarjeta; los puentes H tienen aplicación donde se requiere la inversión de giro de motores de CD, gobernados por un microcontrolador (PICAXE 18X).

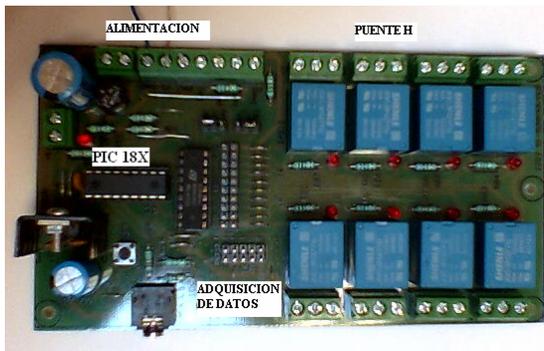


Fig. 8 Tarjeta electrónica de excitación y control.

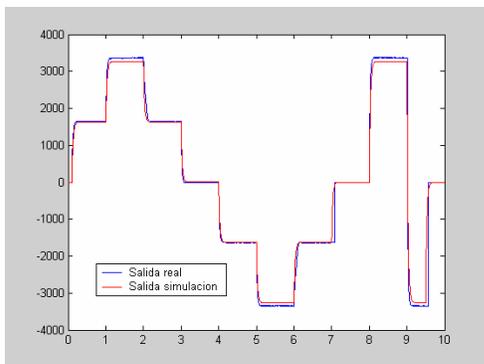


Figura 9. Salidas (rpm) real y simulada a partir del modelo.

El microcontrolador fue se le cargo un programa que coordina los relevadores a fin de tener la secuencia correcta para comando de la silla.

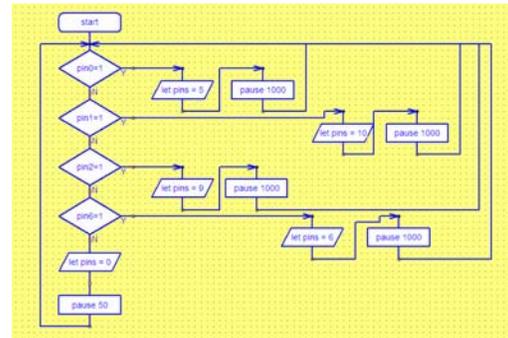


Fig. 10 Programa para controlar los motores.

CONCLUSIÓN

El desarrollo de este proyecto es una importante ayuda para la comunidad con capacidades diferentes en el país.

Sin duda alguna, la tecnología avanza a pasos agigantados y el proyecto es una aportación en la rama de interfaces hombre-máquina en el que la voz juega un papel preponderante para su aplicación.

La mecatrónica ha logrado importantes avances con implementaciones de este tipo, donde investigadores y estudiantes explotan sus capacidades para alcanzar una finalidad totalmente práctica.

CRONOGRAMA

ACTIVIDAD	MAYO				JUNIO				JULIO				AGOSTO				SEP.				OCT.				NOV.				DIC.			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Investigación de reconocimiento de voz	■																															
Metodología para el desarrollo de software			■																													
Desarrollo del software					■	■	■	■	■	■	■	■																				
Sistema de transmisión													■	■	■	■																
Circuito de potencia																	■	■	■	■												
Implementación del control																					■	■	■	■								
Conjunción de sistemas																									■	■	■	■				
Detección de fallas																													■	■	■	■
Optimización del prototipo																																■

REFERENCIAS

- [1] de Pérez T. "Guía de Teoría Señales y Sistemas". U.S.B. 1994
- [2] Kadambe S. and Boudreaux-Bartels G.F., "Linear and quadratic time-frequency signal representations", IEEE Transactions on Signal Processing, vol. 40, N° 10, pp. 2498-2517 oct. 92,
- [3] Daubechies I. "The wavelet transform, time-frequency localization and signal analysis", IEEE Transactions on Information Theory, vol. 36, N° 5, September 1990.
- [4] Klema V. and Laub A. "The singular value decomposition: its computation and some applications," IEEE Trans. on Automatic Control, AC-25(2) pp. 164-176, April 1980.
- [5] Lawson L. and Hanson J. "Solving Least Squares Problems," Prentice-Hall, Englewood Cliffs, N. J., 1974.
- [6] Stewart W. "Introduction to Matrix Computations," Academic Press, Orlando Florida, 1973.
- [7] Rossignol S. and Rodet X. "Features extraction and temporal segmentation of acoustic signals". Proc. ICMC 98, pages 199-202, 1998.
- [8] Molgedey L. and H. Schuster G.. "Separation of a mixture of independent signals using time delayed correlations". Phy. Rev. Lett., 72(23):3634-3637, 1994.
- [9] Oppenheim A. and Schaffer R.; "Discrete-Time Signal Processing", Prentice-Hall, USA. 1989.
- [10] Manual "LabVIEW™ Help", August 2005