

# Sistema de comando remoto por voz para un móvil

Hugo Velasco  
Miguel Escandón  
César Rodríguez

## RESUMEN

En este documento se presenta un estudio progresivo de algunas de las principales tecnologías disponibles para el reconocimiento de voz y comunicaciones inalámbricas. Finalizando con la implementación de un sistema móvil, que responde a la voz de un usuario, haciendo uso de métodos eficientes para el control de movimiento.

El sistema de reconocimiento de comandos de voz conseguido tras la evaluación de varias opciones, se logró a través de la utilización del circuito integrado HM2007, en el cual se encuentra implementada una red neuronal para la identificación de comandos vocales. Los comandos son codificados y transmitidos posteriormente vía infrarrojo. La información llega al móvil y es procesada por los sistemas de control digital, que gobiernan de manera directa su comportamiento. El sistema de reconocimiento de voz es portátil e independiente del móvil, al que comanda de forma remota, de modo que se puede adaptar a otras tareas, que impliquen el control por voz.

**Palabras clave:** Reconocimiento de comandos de voz, redes neuronales, comunicación IR, manejo eficiente de la energía.

## ABSTRACT

This document presents a study of some of the main technologies available for speech recognition and wireless communications. Also the implementation of a mobile system is presented, this mobile is able to respond to the user's voice, by using of efficient methods for the movement control.

The speech recognition system of voice commands obtained after evaluation of several options, was achieved through the use of integrated circuit HM2007, which is implemented in a neuronal network, for vocal commands identification. The commands are codified and transmitted later by infrared way. The information arrives at the mobile and is processed by the control digital systems, which govern of direct way his behavior. The speech recognition system is portable and independent of mobile, which it commands of remote form to, so that it could be adapted to other tasks, that imply the speech control .

## I. INTRODUCCIÓN

En el futuro cercano, será cada vez mas común encontrar sistemas electrónicos como: juguetes,

herramientas eléctricas, computadores y robots, dotados con la capacidad de atender a comandos de voz, dados por el operario del dispositivo. Es por ello que en la actualidad, muchas empresas desarrolladoras de tecnología se encuentran trabajando en prototipos de bajo costo, para brindar al usuario final la facilidad de interfaces hombre-máquina mucho mas amigables, abriéndose de este modo un mercado potencial de productos electrónicos. Estas nuevas tecnologías también contribuirán a mejorar la calidad de vida de personas con discapacidades físicas [1].

Por lo tanto es necesario, hacer uso de la tecnología disponible para llegar a aplicaciones útiles, y de esta manera poder valorar los alcances reales de la misma; evaluando parámetros de confiabilidad, relación costo beneficio y escala de integración entre otros.

Este artículo se organiza de la siguiente manera: en la sección 2 se da a conocer el problema. En la sección 3, se realiza un estudio sobre algunas de las principales tecnologías disponibles tanto para el reconocimiento de voz, como para la transmisión de datos. En la sección 4, se detalla la forma en que fue implementado un modelo de aplicación, para la comprobación del sistema de voz. En la sección 5, se plantea el uso de otra tecnologías para la mejora de los resultados obtenidos.

## II. PLANTEAMIENTO DEL PROBLEMA

Se busca diseñar un sistema de comando remoto y portátil capaz de responder a ordenes dadas por voz humana, y a partir de dichas ordenes, realizar el control de un móvil accionado por motores. El sistema de reconocimiento de voz debe ser adaptable a móviles de cualquier tamaño, para ampliar el rango de aplicaciones (sillas de ruedas, montacargas, juguetes, etc.).

Para llevar a cabo la realización de este sistema se recurrió al uso y contraste de diferentes tecnologías como procesamiento digital de señales (DSP) y sistemas expertos para el reconocimiento del habla, evaluando características relativas a software como tiempo de respuesta y efectividad de interpretación, así como las relativas a hardware considerando consumo de energía y portabilidad.

De igual manera para el desarrollo del sistema de comunicación remota se estudiaron tecnologías de comunicaciones inalámbricas de RF (Radio Frequency) e IR (Infra Red), realizando una comparación de costo / beneficio, eficiencia, y facilidad de manejo, de tal forma que se obtuviera un sistema confiable, y adaptable según las necesidades de la aplicación.

Se busca diseñar un sistema de comando remoto y portátil capaz de responder a ordenes dadas por voz humana, y según estas, realizar el control de un móvil accionado por motores.

### III. DESCRIPCIÓN Y ANÁLISIS DE TECNOLOGÍAS

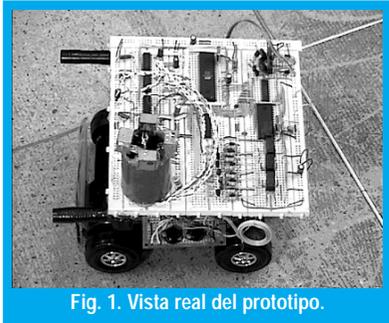


Fig. 1. Vista real del prototipo.

Una vista final del móvil se observa en la figura 1, donde se aprecian los módulos de comunicaciones, potencia, sensorica y control.

La figura 2, presenta el diagrama de bloques del sistema final. Las tres divisiones que se observan allí, corresponden a los siguientes tópicos:

- Sistemas de reconocimiento de voz
- Comunicaciones.
- Sistemas de locomoción y control del móvil.

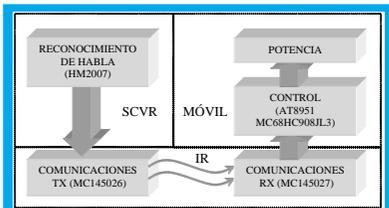


Fig. 2 Diagrama del bloques del sistema de comando remoto. Se muestran las relaciones entre el SCVR, las comunicaciones y el móvil

#### 3.1 Sistema de reconocimiento de comandos de voz

El sistema de reconocimiento de comandos de voz SRCV, es la parte más importante de este proyecto, de su eficiencia para reconocer los comandos de voz, depende en gran parte los resultados finales del sistema.

Para tal efecto, se presentaron dos opciones como las más viables para dar solución al problema. Una primera opción consiste en lograr la identificación de las palabras haciendo uso de análisis estadístico. Así se desarrollo un programa en LabVIEW® [2], sobre una plataforma PC con procesador AMD K6-II de 400MHz y 64MB de memoria RAM. El método implementa el estudio de patrones estadísticos de la voz, tales como: varianza, desviación estándar, tercer momento, etc., y factores deducibles a partir de sus componentes en frecuencia como potencia media, picos de potencia, armónicos etc. Con este primer método se logró una efectividad de aproximadamente un 80% en diferenciar palabras, cuando el programa es configurado para un usuario en particular. Un detalle del software en LabVIEW®, se muestra en la fig. 3.

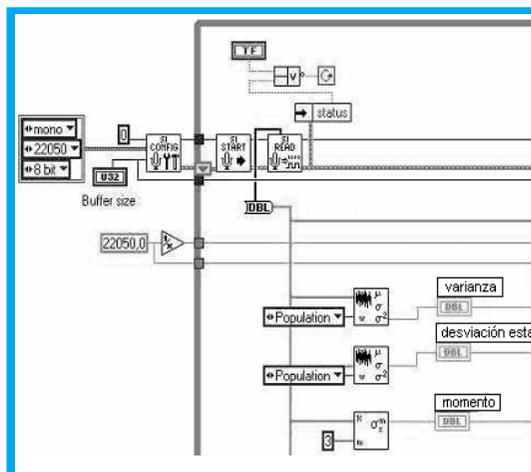


Fig. 3. Detalle del código fuente del software para la identificación de comandos de voz en LabVIEW.

Este método presenta varios inconvenientes debido a que se toman muestras de 0.5 segundos para su análisis. Si este tiempo no coincide con el momento en que se habla, la evaluación de la palabra es incorrecta; además presenta un gran retardo en lograr la identificación del comando dado.

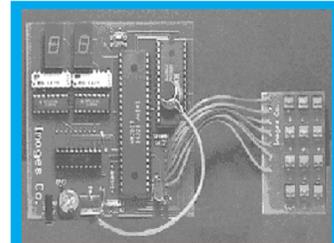


Fig. 4. Presentación final del kit de desarrollo del HM 2007. Se destaca el teclado de programación, el micrófono para la captura de voz y la salida en BCD.

La segunda opción que se estudio fue el integrado HM2007 de H U A L O N Microelectronic Corp., el cual consiste en una red neuronal aplicada al reconocimiento de comandos de voz. En la figura 4 se

muestra el módulo de reconocimiento de voz, a través del cual se toma la señal de voz, se procesa y finalmente, se genera un código digital el cual está asociado al comando que ha reconocido.

Algunas de las características del HM 2007 son:

- Reconocimiento máximo de 40 palabras de 0.96 segundos de duración.
- Reconocimiento mínimo de 20 palabras de 1.92 segundos de duración.
- Soporte de RAM externa, donde se almacenan los descriptores de cada palabra.
- Soporte de micrófono externo.
- Utiliza tecnología CMOS LSI, con lo que se tiene un bajo consumo de energía.

Se escogió este CI debido a su relación costo/pres-taciones, dentro de una moderada gama de integra-dos que se pueden conseguir en el mercado. Este integrado reconoce un 100% más de palabras, que otros que realizan la misma función.

Este integrado facilita enormemente la labor de en-trenamiento y reconocimiento de voz, ya que cuenta con interfaces para teclado matricial 4x3, junto con una salida de datos en BCD, lo cual facilita el manejo y la decodificación de las instrucciones.

Con la red neuronal, se obtuvo una efectividad del 89% para reconocer los comandos entrenados, para la persona que realizó el entrenamiento. Para otras personas varía entre un 60% a un 85% de efectividad.

Después de haber evaluado estas tecnologías para el reconocimiento de voz, se escogió usar el CI HM2007, por cumplir mejor con los requerimientos de portabilidad y relación costo-eficiencia.

#### 3.2. COMUNICACIONES

Cuando se reconoce una palabra previamente en-trenada se le envía un código al móvil para que adopte una opción de movimiento, las palabras que se eligieron se observan en la tabla I junto con el respec-tivo código que se le envía al móvil:

**TABLA I. Voces entrenadas por el usuario y códigos asociados para interpretación por parte del control del sistema, para cada acción.**

ACCIÓN	VOZ DE MANDO	CODIGO	
Adelante	"GO"		00100000
		00110000	01000000
Atrás	"BACK"		00010001
		00100001	00110001
Izquierda	"IZQ"		00010010
		00100010	00110010
Derecha	"DER"		00010011
		00100011	00110011
Alto	"STOP"	00000100	00010100
		00100100	00110100
Acelere	"FAST"	00001000	00011000
		00101000	00111000
Desacelere	"SLOW"	00001001	00011001
		00101001	00111001
Ven	"COME"	00010000	00000001
		00000010	00000011

Es importante que la comunicación entre el móvil y el sistema de voz sea confiable, de manera que se eliminen posibles fallas del sistema. También es necesario tener en cuenta, la posibilidad de una futura ampliación de los sistemas u opciones en el comportamiento del móvil.

El desarrollo del sistema de comunicaciones comenzó con el estudio de factibilidad de un sistema inalámbrico RF basado en el CI TRF6900 de Texas Instruments [3]. Este sistema presentaba bastantes complejidades técnicas para su desarrollo debido a costos, el tiempo de implementación y tolerancia

de los componentes externos, aunque es una opción para tener en cuenta en un trabajo futuro debido a sus características de alcance y disponibilidad del enlace SRCV - Móvil.

Una segunda alternativa para lograr la comunicación entre el SRCV y el móvil utiliza un par de transmisión-recepción de datos fabricado por Motorola™, especial para comunicaciones seriales por infrarrojo, MC145026 (emisor) y MC145027 (receptor). Estos permiten transmitir las palabras ya codificadas hacia el módulo de control aprovechando los datos de un nibble de longitud que se pueden enviar por medio de estos integrados [4], [5].

Dada la confiabilidad en la transmisión lograda, gracias a los sistemas de prevención de errores internos que poseen los integrados MC1450CXX y a su flexibilidad para establecer diferentes canales de comunicación, se adoptó la opción IR, como sistema a implementar en el prototipo final.

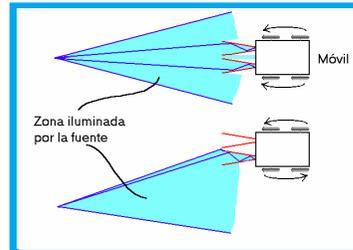
En el prototipo realizado, el operario del móvil cuenta con un transmisor que tiene una dirección que selecciona un canal específico; en el móvil se instala un receptor con el mismo canal habilitado. El circuito de control sabe cual es el respectivo número que identifica a cada una de las instrucciones (ver tabla I) y dependiendo de cual sea este, ejecutará la acción correspondiente al comando de voz.

Además, se utilizaron estos CI para implementar la función de búsqueda de la base remota (el lugar en que se encuentra ubicado el SRCV) por parte del móvil. Esta función se basa en el uso de dos receptores MC145027, aislados de tal forma, que reconocen la fuente de transmisión de manera unidireccional y por medio de un algoritmo de selección), permiten al móvil dirigirse hacia la base de transmisión.

El algoritmo de selección consiste en la evaluación de las señales provenientes de cada uno de los sensores IR. Estos generan una señal en estado alto TTL, en el momento de detectar la fuente de transmisión, que

emite información en el mismo canal de comunicación, seleccionado por el receptor, es decir, el móvil.

Cuando la señal es activa en ambos receptores, el sistema de control toma la decisión de dar marcha hacia adelante a ambos motores, para dirigir el móvil hacia la fuente. Si alguna señal se pierde en este proceso, el control produce un giro correctivo del móvil hacia la dirección en la cual aún se encuentra presente la señal de recepción. Esto se muestra en la figura 5.



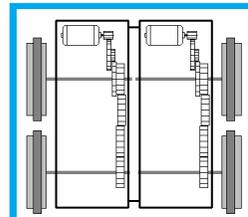
**Figura 5. Comportamiento gráfico del algoritmo de selección. Las flechas indican el sentido de giro de las llantas.**

## IV. MÓVIL

El proceso de diseño tuvo en cuenta las etapas a continuación, en las cuales se fija como parámetro principal la eficiencia del sistema.

### 4.1 Sistema de tracción del móvil

Dado que el móvil debe salir de dificultades cuando se encuentre en condiciones adversas como un terreno escabroso u obstáculos, es importante que el sistema de tracción sea lo mas potente posible.



**Fig. 6. Esquemático del sistema de tracción del móvil.**

Para tal fin, se adoptó como plataforma mecánica del móvil, dos carros de baterías comúnmente conocidos como "Buggy" acoplados lateralmente; un dibujo esquemático del acople se ve en la figura 6.

Cada uno de estos carros cuenta con un motor DC con una tensión nominal de 3 voltios, y caja de transmisión con una relación de 58,57:1. Esto muestra que la velocidad sobre las llantas no es muy alta, pero el torque que se desarrolla en ellas sí lo es. Por lo cual, una carga hasta de una libra puede ser transportada con facilidad (en este caso, la carga está constituida por los sistemas electrónicos y una batería).

Debido a que es necesario controlar el sentido de giro del móvil se hacen girar los motores en direcciones opuestas, lo que da como resultado un movimiento circular alrededor del centro del eje.

### 4.2 Sistema de potencia

Dado que en un sistema móvil, la energía que se puede transportar es limitada, la alimentación puede llegar a ser un problema. La primera solución que se planteó fue la utilización de dos baterías, ya que en la práctica se presentaron problemas de ruido, que afectaban el comportamiento del circuito de control. Lo

Debido a que es necesario controlar el sentido de giro del móvil se hacen girar los motores en direcciones opuestas, lo que da como resultado un movimiento circular alrededor del centro del eje.

anterior en el momento en que se efectuaba algún cambio de estado sobre los motores, por lo tanto fue necesario desacoplar totalmente el sistema de alimentación, de los sistemas de tracción y control.

La alimentación del móvil es una batería de 6V @ 4Ah (Amperios hora). Fue escogida porque brinda una autonomía bastante amplia para el funcionamiento de los sistemas. La batería tiene unas dimensiones (108mm x 70mm x 47mm) y peso (0.9kg), apropiados para la capacidad de carga del sistema de tracción. Para obtener la alimentación regulada del circuito de control, se utilizó un conversor reductor LM2678T-5 de National Semiconductors [6], de 5V @ 3A.

El manejo de los motores se realiza por medio de relevos electromecánicos, configurados como puente H, para conseguir la inversión de la polaridad. Se hace necesario además, minimizar las pérdidas que puedan tener los interruptores que se conmutan en el puente H. Los relevos utilizados para la implementación son los ATX221 M01 de doble bobina fabricados por NaiS.

Estos relevos cuentan con memoria, ya que con sólo un pulso de corriente se logra el cambio de estado y sostenimiento de éste, sin necesidad de mantener energizada la bobina, esto por cada interruptor. Con la aplicación de un pulso sobre la otra bobina se logra cambiar el estado del interruptor.

En comparación con los relevos clásicos, que deben mantener la alimentación del electroimán para conservarse en cierto estado, se logra un ahorro significativo de energía. Esto en vista a que los cambios de acción ocurren en un tiempo pequeño, con respecto al tiempo que dura realizando una acción en particular.

Para administrar la energía a los motores se utiliza una modulación por ancho de pulso (PWM) [7], [8], de la fuente de alimentación, con lo que se obtuvo una mejora sustancial en la autonomía del móvil, al compensarse la caída de tensión que se presenta con el tiempo. Esto se logra con un incremento del porcentaje de ciclo útil máximo de la señal, con que se excita a los motores. Este sistema se implemento por medio del módulo de temporización del  $\mu\text{C}$  MC68HC908JL3 de Motorola [9], que funciona como PWM.

#### 4.3 Sistema de control del móvil

El sistema de control del móvil se implementa por medio de un microcontrolador 8951 de ATMEL [10] junto al  $\mu\text{C}$  MC68HC908JL3. La selección de estos dos  $\mu\text{C}$  se debe al aprovechamiento de las características individuales, como número de puertos I/O disponibles en el 8951, y los periféricos de temporización y conversión analógico-digital que posee el JL3. La función que desempeña este conjunto hace posible recibir las señales de los sistemas de comunicación y con base en los códigos recibidos tomar las decisiones correspondientes, como se describió anteriormente.

## V. TRABAJO A FUTURO

Después de la realización de un pequeño modelo del sistema de comando remoto, se podría llevar a

cabo la ejecución de un móvil a mayor escala, que permita transportar una persona.

También se proyecta el uso de tecnologías, como la transformada Wavelet, que permitan obtener un mayor desempeño del sistema, en cuanto a la efectividad del reconocimiento de los comandos. Es decir, hacer posible que cualquier usuario pueda operar el sistema, con un 100% de confiabilidad.

## CONCLUSIONES

Dado que los requerimientos tanto de hardware como de software, necesarios para realizar el procesamiento de las señales y en este caso señales de voz son tan grandes, los tiempos de respuesta que se deben esperar de ningún modo cumplen con exigencias de tiempo real. Por lo tanto las aplicaciones en las cuales se utilice este sistema, no podrán demandar tiempos despreciables de respuesta (tiempos menores a 2 seg).

Cuando se realizan implementaciones, para módulos que requieren determinadas condiciones de autonomía y consumo de energía, la relación peso vs. capacidad Ah es un doble compromiso por parte del diseñador. Es decir, se requiere llegar a un rango óptimo, donde se presente el mínimo peso con la mayor autonomía.

De acuerdo con las pruebas realizadas al prototipo final, durante el día, en un espacio cerrado e iluminado con lámparas fluorescentes, se determinó que se pueden alcanzar distancias máximas de transmisión de 15 m.

La interferencia de la luz es un problema crítico en el proceso de comunicaciones IR, entre los elementos del sistema, ya que afecta los umbrales de detección de luz infrarroja.

El sistema cuenta con una efectividad para responder a los comandos de 60% a 85%, para cualquier usuario.

## REFERENCIAS

- [1] SEÑAL COLOMBIA. (Documental en Video). Bogotá. Diciembre 2001.
- [2] NATIONAL INSTRUMENT. Labview basics course manual. April 1998 Edition.
- [3] Sprevak P., y Loy M. Texas Instruments. Application Report. SWRA034B - May 2001.
- [4] MOTOROLA. Encoder and Decoder Pairs. MC145026/D. Revisión 3. 1998.
- [5] JAN AXELSON. Comunicaciones inalámbricas con PC. Electrónica y computadores. Cedit. 1997
- [6] NATIONAL SEMICONDUCTOR. Power Ics Databook. USA. 1995.
- [7] SAVANT, RODEN y CARPENTIER, Diseño Electrónico. Addison-Wesley Iberoamericana. USA. 1992.
- [8] BOYLESTAD, Robert L.; NASHELKY, Louis. ELECTRONICA: Teoría de Circuitos. 6ª Ed. Prentice Hall Hispanoamericana S.A., México DF. 1996.
- [9] MOTOROLA. MC68HC908GP32/H Technical Data. Revisión 5. Julio 2001.
- [10] BERNAND ODANT. Microcontroladores 8051 y 8052. 1ª Ed. Editorial Paraninfo. Madrid. 1995.

### Hugo Fernando Velasco Peña

Estudiante de Ingeniería Electrónica, Universidad Distrital. Integrante del grupo de Instrumentación Virtual. hvelasco@ieee.org

### Miguel A. Escandón Moscoso

Estudiante de Ingeniería Electrónica, Universidad Distrital. Integrante del grupo de Instrumentación Virtual. miguelescandon@mixmail.com

### César H. Rodríguez Garavito

Estudiante de Ingeniería Electrónica, Universidad Distrital. Integrante del grupo de Instrumentación Virtual. cesar\_rodriguez80@yahoo.es