

Control de desplazamiento de móvil robótico mediante movimientos de cabeza utilizando visión de máquina

Displacement control of mobile robot through head movements using machine vision

R. Jiménez y O. F. Avilés

Recibido Noviembre 1 de 2013 – Aceptado Junio 6 de 2014

Resumen - En este artículo se presenta el control de movimiento de un móvil robótico de cuatro ruedas con control de dirección, el cual es operado mediante comandos basados en movimientos de cabeza para su desplazamiento en avance, reversa, izquierda y derecha. Mediante técnicas de procesamiento de imagen se genera un vector de desplazamiento de cabeza, cuya dirección es utilizada para determinar la orientación de desplazamiento del móvil. Los movimientos de cabeza se basan en la detección de los ojos mediante el algoritmo de Viola-Jones, de este se determina una posición de referencia a fin de establecer las coordenadas de origen del vector, de forma tal que el desplazamiento de ojos desde este punto, genera la magnitud que determina la dirección de desplazamiento del móvil, el telecontrol se realiza mediante protocolo XBEE del computador a la tarjeta microcontrolada que opera el vehículo, la precisión obtenida bajo un ambiente de luz semi controlado es de 99.8%.

Palabras Clave: tele control, visión 2D, clasificador haar, procesamiento de imagen, detección de ojos.

Abstract - In this article the motion control of a robotic mobile-wheel steering control, which is operated by head movements based for movement forward, reverse, left and right commands is presented. By image processing techniques displacement vector head, whose address is used to determine the direction of movement of the mobile is generated. Head movement based on the detection of the eye by the algorithm of Viola-Jones of this a reference position to establish the coordinates of the origin of the vector so determined such that the movement of eyes from this point generates the magnitude which determines the direction of movement of the mobile, remote protocol is performed by the computer XBEE microcontrolled card to operating the vehicle, the accuracy obtained under semi controlled lighting environment is 99.8%.

Key Words: tele control 2D vision, haar classifier, image processing, eye detection.

1. INTRODUCCIÓN

Los sistemas de visión de máquina, soportados por algoritmos de procesamiento de imagen, se emplean hoy día en múltiples facetas del área de ingeniería. Sus aplicaciones van desde dispositivos de seguridad, automatización, control robótico y medicina. Presentándose así como una poderosa herramienta de desarrollo y aplicaciones, que ofrecen muchos tópicos aún por abordar en el campo de la investigación.

En el campo de seguridad se encuentran trabajos referentes a detección de rasgos característicos que permiten identificar

¹Producto derivado del proyecto de investigación “Control de un móvil robótico por medio de movimientos de cabeza”, apoyado por la Facultad de Ingeniería, Programa de Ingeniería Mecatrónica, Universidad Militar Nueva Granada a través del grupo de investigación en Aplicaciones Virtuales GAV.

R. Jiménez y O. F. Avilés, son docentes de la Universidad Militar Nueva Granada, Bogotá, Colombia, (correos e.: robinson.jimenez@unimilitar.edu.co, oscar.aviles@unimilitar.edu.co).

una persona para restringir o controlar el acceso a ciertas áreas [1]. Se presentan desarrollos en el campo de detección de estados de riesgo de un conductor frente al volante, como en detectar estados del ojo que indiquen somnolencia, por ejemplo [2][3]. En [4] se presenta un sistema de control de tráfico vehicular detectando la cantidad de vehículos mediante una cámara en la vía.

En el campo de la automatización se presentan desarrollos para el reemplazo de sensores, como el presentado en [5], en el cual se utilizan técnicas de procesamiento de imagen para distinguir cada uno de los tres tanques de la escena y realizar la medida de nivel de llenado de cada uno. En el campo de control robótico se tienen aplicaciones de obstáculos para evasión y/o cálculo de trayectorias [6] [7]. En el campo de la medicina se encuentran muchas aplicaciones, varias de ellas encaminadas a detección de anomalías, calcificaciones o tumores en imágenes ultrasónicas o radiológicas [8].

El presente artículo expone la aplicación de un algoritmo de visión de máquina para el telecontrol de un móvil robótico mediante algoritmos de visión de máquina. Para el caso se realizan pruebas con un pequeño vehículo robótico, pero se puede aplicar a brazos robóticos o cualquier otra plataforma con al menos dos grados de libertad. Entre los casos prácticos posibles, se prevé su uso para ayuda a pacientes con discapacidad motora, orientación de fuentes de luz en el caso de un cirujano al operar o incluso el control del ángulo de visión del espectador en un ambiente virtual.

El telecontrol aplicado se soporta por medio de un módulo de comunicación inalámbrico XBEE, que transmite las señales procesadas con la cámara de adquisición desde el equipo de cómputo (PC) hasta la tarjeta microcontrolada que opera el móvil, el cual es capaz de desplazarse en avance, retroceso, izquierda o derecha.

En adelante el documento está organizado de la siguiente forma: la Sección 2 describe la técnica de detección de movimientos de cabeza, en la Sección 3 se presenta el análisis de resultados y finalmente en la Sección 4 las conclusiones.

II. DETECCIÓN DEL MOVIMIENTO DE CABEZA

El sistema de visión de máquina utilizado, está basado en la captura de imágenes desde el computador, por medio de una cámara web interna cuya resolución es de 640 x 480 píxeles, procesadas mediante la librería de Intel OpenCV 2.1, bajo un entorno de programación en C de uso libre denominado Eclipse, versión Galileo.

En este ambiente de programación se implementó el algoritmo de Viola – Jones [9], el cual opera en función a un clasificador AdaBoost [10] entrenado para detección de los ojos. Para el entrenamiento de este, se requiere un amplio grupo de imágenes (al menos 5000), las cuales deben incluir tanto ojos (imagen positiva) como diferentes imágenes que no los incluyan (imágenes negativas) y que permitan una

clara discriminación. Para el desarrollo de este trabajo se emplea el clasificador ya entrenado, que hace parte de las librerías de visión de máquina utilizadas y el cual se ha utilizado en varios trabajos como [11] y [12].

De forma general el clasificador Adaboost es entrenado para la identificación de un objeto determinado, bajo una serie de clasificadores débiles, descritos mediante la ecuación (1) y soportado en descriptores haar.

$$h_j(x) = \begin{cases} 1 & p \cdot f_j(x) > p \cdot \theta_j \\ -1 & \text{en otro caso} \end{cases}$$

En la ecuación (1), x corresponde a la imagen, f es el valor de la j característica Haar aplicada a la imagen, p es la polaridad de la característica relacionada con la orientación y θ es el umbral que decide si la característica evaluada es representativa sobre la imagen.

El clasificador entrenado permite la detección de los ojos indistintamente del sexo, raza o edad. Es robusto a condiciones de iluminación variables cuando las imágenes tienen fuente de luz día directa. La figura 1 ilustra el resultado del clasificador utilizado en diferentes personas, en la parte superior las imágenes se tomaron en horas de la mañana y las de la parte inferior en horas de la tarde.

La tabla 1 relaciona la matriz de confusión obtenida al contabilizar la detección del ojo, realizada por el clasificador, en una serie de videos de, en promedio, 1 minuto de duración para 30 usuarios. Se evidencia la alta sensibilidad y precisión del clasificador, lo que evidencia los buenos resultados del mismo. Algunos de los falsos negativos, los cuales corresponden a casos en donde se encuentra el ojo sin que realmente esté, se deben a condiciones de inclinación lateral de la cabeza y/o reflexión de lentes que no permiten distinguir el ojo claramente.



Fig. 1. Resultado del clasificador

TABLA I. MATRIZ DE CONFUSIÓN DETECCIÓN DEL OJO

Detección de Ojos		PREDECIDO (%)	
		POSITIVO	NEGATIVO
REAL	FALSO	31	46
	VERDADERO	15830	215
Sensibilidad		99,71	
Especificidad		87,40	
Precisión		99,80	

III. MOVIMIENTO DEL MÓVIL

Para establecer la dirección de movimiento de la cabeza en función de la detección de los ojos realizada mediante el clasificador Haar, se obtiene un vector de dirección cuyo punto de origen es calculado al tomar el centro de los rectángulos que delimitan la región de cada ojo, cada centro es denominado O_{Ixy} y O_{Dxy} , respectivamente. Se calcula la distancia en línea recta entre ambos centros, dividida entre dos, este punto es tomado como referencia o coordenada inicial $C_i(0,0)$; tal como se ilustra en la ecuación 4. Para asegurar un punto fijo de esta esta coordenada, se toma el promedio de los 15 primeros frames de la señal de video obtenidos de la cámara y a este nuevo punto origen se le denomina $C_{ip}(0,0)$, la ecuación 5 ilustra el proceso de cálculo de este. Después del quinceavo frame, el desplazamiento del punto medio entre los ojos es tomado como posición final o extremo del vector de dirección, estableciendo así la dirección de desplazamiento, como se ilustra en la figura 2, donde el punto final se destaca en amarillo.

$$C_i(0,0) = (\sqrt{O_{Ixy}^2 - O_{Dxy}^2})/2 \quad (4)$$

$$C_{ip}(0,0) = \frac{\sum_{n=0}^{14} C_i(0,0)_n}{15} \quad (5)$$

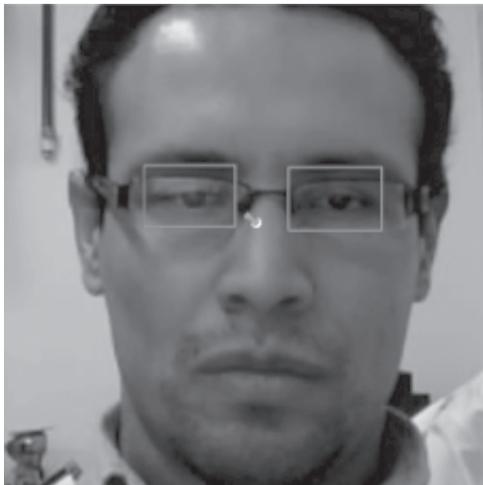


Fig. 2. Vector resultante

La figura 3 ilustra el resultado gráfico del vector obtenido al realizar los diferentes movimientos de la cabeza que permitirán al móvil desplazarse, en orden de aparición se tiene, movimiento hacia abajo, movimiento hacia la izquierda, movimiento hacia la derecha y movimiento hacia arriba.



Fig. 3. Desplazamientos obtenidos

Las pruebas de validación para determinar la efectividad del vector de dirección, consistieron en evaluar el algoritmo de movimiento de cabeza con cinco usuarios diferentes a los empleados para pruebas del algoritmo de clasificación para detección de ojos. Para ello se replicaron los movimientos ilustrados en la figura 3, cinco veces por cada usuario, obteniendo los resultados observados en la tabla 2. Se puede apreciar que en general la detección del movimiento deseado es detectada acertadamente por el algoritmo sin importar el usuario, que para el caso fueron tres hombres y dos mujeres.

Los errores detectados obedecen a movimientos extremos de la cabeza hacia un lado sin retornar nuevamente al punto de referencia, donde se pierde la detección de uno de los ojos, de forma tal que como resultado no se registra un nuevo movimiento hacia el lado o hacia arriba, razón por la cual no es contabilizado.

TABLA II. MOVIMIENTOS REGISTRADOS

Usuario	Arriba	Abajo	Izquierda	Derecha
1	5/5	5/5	5/5	5/5
2	4/5	5/5	5/5	5/5
3	5/5	5/5	4/5	5/5
4	5/5	5/5	5/5	5/5
5	5/5	4/5	4/5	5/5

Una vez detectada la dirección de movimiento de la cabeza, esta se envía, codificada por comunicación serial, a un puerto del computador, según se indica en la tabla 3.

TABLA III. CODIFICACIÓN DE MOVIMIENTOS

Movimiento	Código
Izquierda	1
Derecha	2
Arriba	3
Abajo	4

Bajo esta codificación en el móvil se corre el algoritmo de control indicado en el diagrama de flujo de la figura 5.

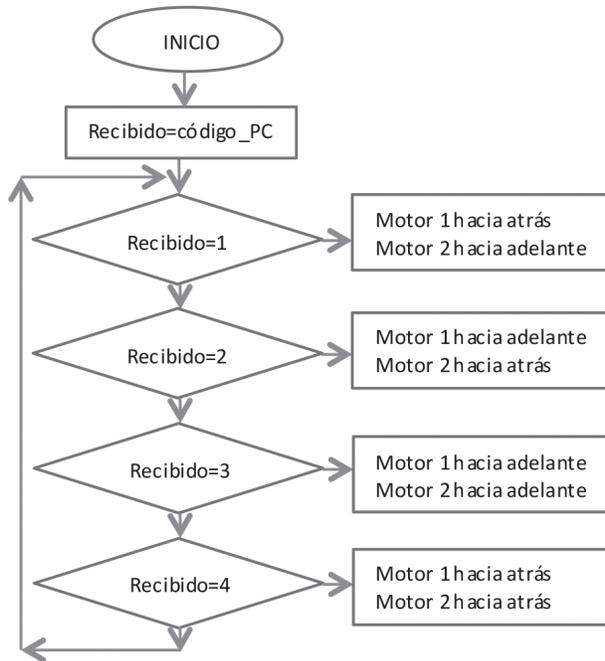


Fig. 4. Diagrama de flujo del móvil

La codificación obtenida de la detección del movimiento de la cabeza que asocia la dirección de desplazamiento para el móvil es entregada por comunicación serial desde el entorno de C a un módulo transmisor XBEE conectado al puerto del computador. Este retransmite la señal de control a un módulo receptor en el móvil, conectado a una tarjeta microcontrolada encargada de generar la salida de control a un circuito integrado de dos puentes H de referencia L293B, uno para la dirección izquierda-derecha y otro para avance o retroceso.

El módulo de radiofrecuencia corresponde a la referencia XBee Pro, el cual cumple con el estándar IEEE 802.15.4, y cuenta con una potencia de transmisión de 60 mW, una sensibilidad de recepción de -100 dbm y un alcance de hasta 1 KM en espacios abiertos. A fin de disminuir el riesgo de interferencias externas, como pueden ser las redes WiFi, se emplea en el canal 12 del mismo, la comunicación es unidireccional y llega a otro módulo Xbee igual, que por comunicación serial tipo RS232, entrega la información de desplazamiento al microcontrolador. Los retardos obtenidos en la comunicación no son significativos y se presentan en

el orden de máximo centenas de milisegundo operando el móvil correspondientemente a la dirección del movimiento del usuario. La figura 4 ilustra el móvil empleado.

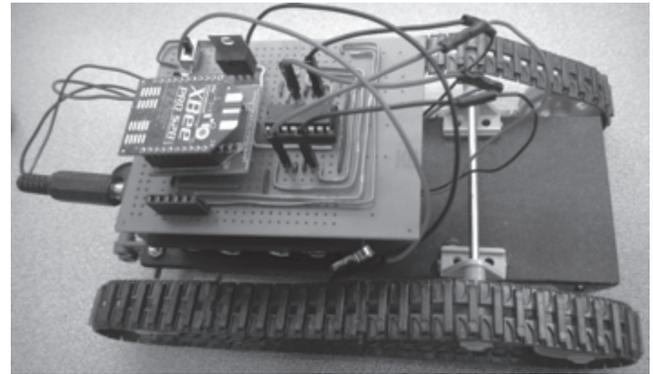


Fig. 4. Móvil empleado

Debido a que el control del móvil se realiza una vez determinado el movimiento de cabeza, sólo se envían transmisiones cuando se detecta un cambio en la dirección o sentido movimiento, por lo que las pérdidas presentadas en la tabla 2 no afectan el desempeño en el control del mismo.

La característica de que el móvil sea tipo oruga, permite que se desplace más fácilmente en diferentes ambientes. El sistema responde a velocidades de movimiento de cabeza normales, trabajando a 30 frames por segundo, debido a que la detección del algoritmo de clasificación está en promedio a los 15 milisegundos.

IV. ANÁLISIS DE RESULTADOS

El equipo de cómputo en el que se implementó el sistema de detección del movimiento de cabeza, cuenta con un Procesador de 2.4GHz y 2 GB de memoria RAM, lo que da a la aplicación la capacidad de operar eficientemente en tiempo real, es decir a 30 frames por segundo que es la velocidad de toma de imágenes de la cámara de video utilizada.

Se establecieron dos escenarios de prueba con los cinco usuarios relacionados en la tabla 2. El trayecto uno consiste en un recorrido rectilíneo, mientras que el trayecto dos en un ocho invertido, tal como se ilustra en la figura 5.

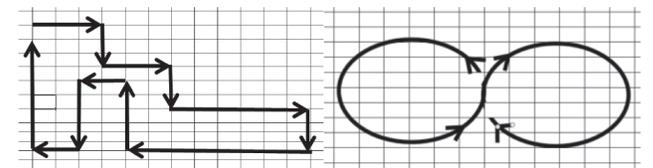


Fig. 5. Trayectorias empleadas

Los recorridos fueron implementados con líneas guía en el suelo. Se evaluó el tiempo que le tomo a cada usuario completar el recorrido y la precisión del mismo, la cual se calculó como el promedio de la desviación del móvil respecto a la guía, en relación a cada uno de los recorridos.

Los resultados obtenidos se presentan en la tabla 4.

TABLA IV. MOVIMIENTOS DE CONTROL DEL MÓVIL

	Usuario	1	2	3	4	5
TRAYECTO 1	Tiempo (seg)	25	23	20	25	26
	Precisión (%)	95	92	94	91	95
TRAYECTO 2	Tiempo (seg)	52	54	59	50	52
	Precisión (%)	79	78	76	80	76

Se puede apreciar, en la tabla 4, que mientras más compleja la ruta más difícil es tener precisión en el seguimiento de la misma y por consiguiente mayor tiempo tomará. Se evidencia igualmente que el algoritmo implementado no incide significativamente en el control del móvil al cambiar de usuario, dado que se tienen tiempos y precisiones muy similares para cada recorrido.

Finalmente, cada usuario logro realizar todo el recorrido deseado, lo cual evidencia la facilidad de uso del sistema de control mediante el algoritmo de visión de máquina empleando el clasificador y su versatilidad al momento de comandar un móvil.

V. CONCLUSIONES

Se logró implementar un algoritmo de detección de ojos altamente eficiente y con mínimas variaciones bajo los cambios de luz día al tener una visión frontal del rostro. Dicho desempeño baja al presentarse sombras y movimientos de cabeza principalmente hacia arriba, llegando incluso a perder la referencia, por lo cual los movimientos hacia arriba deben ser moderados.

Los movimientos de cabeza establecidos son eficientemente detectados cuando parten desde el punto de referencia, si se presentan movimientos extremos y no se retorna a dicha referencia, los nuevos movimientos hacia el mismo sentido no se contarán, lo cual dependiendo de la aplicación puede ser inconveniente.

En conjunto los algoritmos implementados permitieron un control del móvil a voluntad del usuario independientemente de este y sin retardos en la ejecución de los comandos de direccionamiento.

Como trabajo futuro se logra determinar que la magnitud del movimiento de cabeza se puede relacionar con cambios de la velocidad de desplazamiento del móvil.

Las aplicaciones que se puedan derivar del proyecto están orientadas a casos como comando de sillas de ruedas por personas con limitaciones físicas.

AGRADECIMIENTOS

Los autores agradecen a la Vicerrectoría de

Investigaciones de la Universidad Militar Nueva Granada, la cual apoyo financieramente el desarrollo del proyecto, mediante el ING 1270, el cual fue desarrollado al interior del grupo de investigación GAV del programa de ingeniería en Mecatrónica.

REFERENCIAS

- [1] Jain, A.K., "Biometrics: Proving Ground for Image and Pattern Recognition," Image and Graphics, 2007. ICIG 2007. Fourth International Conference on, vol., no., pp.3,3, 22-24 Aug. 2007.
- [2] David Crundall, Geoffrey Underwood, Chapter 11 - Visual Attention While Driving: Measures of Eye Movements Used in Driving Research, In: Bryan E. Porter, Editor(s), Handbook of Traffic Psychology, Academic Press, San Diego, 2011, Pages 137-148, ISBN 9780123819840, 10.1016/B978-0-12-381984-0.10011-6.
- [3] Robinson Jiménez, Flavio Prieto y Victor Grisales. "Detection of the tiredness level of drivers using machine vision techniques", CERMA 2011 IEEE. Pg 97-102.
- [4] Espinosa Fabio, Aviles Oscar, Gordillo Camilo y Robinson Jiménez. Machine Vision algorithms applied to dynamic traffic light control. Revista DYNA V. 80 Fasc 180, pg 56, 2013. Issn 0012-7353.
- [5] Robinson Jimenez Moreno, Oscar Aviles Sanchez, Fabio Espinosa, "Level measurement comparison between 3D vision system based on Kinect and ultrasonic industrial sensor". Revista Asian Transactions On Engineering ISSN: 2221-4267 ed: v.2 fasc.5 p.10 - 19 ,2012.
- [6] Restrepo, J.C.; Villegas, J.; Arias, A.; Serna, S.; Madrigal, C., "Trajectory generation for a robotic in a robocup test scenery using Kalman filter and B-spline curves," Image, Signal Processing, and Artificial Vision (STSIVA), 2012 XVII Symposium of , vol., no., pp.110,115, 12-14 Sept. 2012.
- [7] Ruiz Hidalgo, D.A.; Velasco Hernandez, G.A.; Calvache Vallejo, B.A.; Caicedo Bravo, E.F., "Path generation and mapping based on global vision for mobile robots in structured environments," Image, Signal Processing, and Artificial Vision (STSIVA), 2012 XVII Symposium of , vol., no., pp.168,176, 12-14 Sept. 2012.
- [8] Weiyang Xie; Yide Ma; Bin Shi; Zhaobin Wang, "Gallstone segmentation and extraction from ultrasound images using level set model," Biosignals and Biorobotics Conference (BRC), 2013 ISSNIP , vol., no., pp.1,6, 18-20 Feb. 2013.
- [9] P. Viola, M. Jones. "Rapid object detection using a boosted cascade of simple features[C]". Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.12, 2001, pp. I-511-I-518.
- [10] Yong Ma; Xiaoqing Ding, "Robust real-time face detection based on cost-sensitive AdaBoost method," Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on , vol.2, no., pp.II,465-8 vol.2, 6-9 July 2003.
- [11] R. Jiménez Moreno, S. A. Orjuela, P. Van Hese, F. A. Prieto, V. H. Grisales and W. Philips. "Video surveillance for monitoring driver's fatigue and distraction", Proc. SPIE 8436, 84360T (2012); <http://dx.doi.org/10.1117/12.922085>.
- [12] Jia Mingxing; Xu Hengyuan; Wang Fei; , "Research on driver's face detection and position method based on image processing," Control and Decision Conference (CCDC), 2012 24th Chinese , vol., no., pp.1954-1959, 23-25 May 2012.

Robinson Jiménez Moreno. Ingeniero Electrónico graduado de la Universidad Distrital Frco. Jose de Caldas, Bogotá 2002. Magister en Ingeniería en el área de Automatización industrial de la Universidad Nacional de Colombia, Sede Bogotá, 2012. Actualmente docente del programa de Ingeniería en Mecatrónica de la Universidad Militar Nueva Granada, Bogotá, Colombia.

Oscar F. Aviles Sanchez. Ingeniero Electrónico graduado de la Universidad Antonio Nariño, Bogotá. Maestría en Sistemas Automáticos de producción en la Universidad Tecnológica de Pereira, Colombia. Doctor de la Universidad Estatal de Campinas, Brasil. Actualmente docente del programa de Ingeniería en Mecatrónica de la Universidad Militar Nueva Granada, Bogotá, Colombia