

January 2002

¿Puede el computador mirar a los ojos?

Sergio A. Rojas

Universidad de La Salle, Bogotá, revista_uls@lasalle.edu.co

Follow this and additional works at: <https://ciencia.lasalle.edu.co/ruls>

Citación recomendada

Rojas, S. A. (2002). ¿Puede el computador mirar a los ojos?. *Revista de la Universidad de La Salle*, (33), 11-19.

This Artículo de Revista is brought to you for free and open access by the Revistas de divulgación at Ciencia Unisalle. It has been accepted for inclusion in *Revista de la Universidad de La Salle* by an authorized editor of Ciencia Unisalle. For more information, please contact ciencia@lasalle.edu.co.

¿PUEDE EL COMPUTADOR MIRAR A LOS OJOS?

Sergio A. Rojas

Ingeniero de Sistemas, Universidad Nacional
Docente Facultad de Ingeniería de Diseño y
Automatización Electrónica
Director Grupo de Interfaces Naturales (GIN)
Universidad de La Salle

Durante décadas hemos sido testigos de la intromisión del computador como herramienta indispensable para el apoyo de la mayoría de labores humanas. Como efecto lateral, ha aparecido una nueva forma de "analfabetismo" que se supera aprendiendo a utilizar la máquina y acomodándonos a su forma de comunicarse. Nos hemos familiarizado con el ratón y la impresora, cuando antes preferíamos una conversación o una carta escrita. Una nueva tendencia en la informática, denominada Interacción humano - máquina (HCI, por sus siglas en inglés) estudia el uso de canales de comunicación, con las máquinas, que sean más familiares para las personas: el habla, la gesticulación o el texto manuscrito. En este artículo se describe un proyecto relacionado con el tema, desarrollado en la Facultad de Ingeniería de Diseño y Automatización Electrónica, y que busca permitir que un usuario sentado frente a un computador pueda transmitirle información a través de la mirada.



Con el advenimiento del computador hemos debido adaptarnos a los dispositivos de entrada y salida como el teclado, el ratón y la pantalla, que son muy apropiados para procesar la información digital que consume la máquina, pero que son poco amigables para nuestra manera de comunicarnos. Claro, para que el procesador pueda desplegar una letra en la pantalla, es mucho más sencilla la acción de oprimir una tecla que genere una señal eléctrica para que se active un banco de memoria, que a su vez recupere una cadena de números binarios que la represente. Para nosotros, sería mucho más práctico y ameno, simplemente pronunciar la letra, con la lejana esperanza de que el computador pudiera "escucharla". Esta lejana esperanza se ha estado convirtiendo en una de las más activas áreas de investigación de los científicos de la computación.

La interacción humano - máquina (HCI) es una ciencia que investiga cómo amenizar la comunicación con el computador, por medio de interfaces que en principio sean más comprensibles y más fáciles de manipular (sonidos, imágenes, gestos, texto manuscrito). Varias empresas y universidades están desarrollando proyectos con resultados muy prometedores en el reconocimiento de textos escritos a mano, la imitación de gestos humanos, la medición de características biométricas para reconocimiento de personas, la detección de emociones con señales del tacto, y otras más. El robot "Cog" desarrollado en MIT^[1], es capaz de seguir un objetivo que la persona tenga en la mano, moviendo su cabeza como un perrito al que le van a dar una galleta; además es capaz de remedar la expresión del rostro de su dueño, acompañándolo en su congoja, o regocijándose con él. En IBM se ha desarrollado el *Emotion Mouse*^[2], un ratón que no solo sirve para guiar el cursor en la pantalla sino que además, mediante las señales de tacto de la mano, puede

identificar el estado de ánimo del usuario para por ejemplo, desplegarle un mensaje de aliento.

Con el ánimo de comenzar un trabajo exploratorio en este tipo de sistemas, el GIN de la Facultad de Ingeniería de Diseño y Automatización Electrónica de la Universidad de La Salle, ha propuesto el proyecto de investigación MIMICO, para estudiar una nueva interfaz de comunicación mediante la mirada: que el computador pueda "ver" al usuario que tiene al frente, determinar hacia dónde está mirando y, entonces, seguir las instrucciones del caso, ya sea que por que esté concentrado en alguna región de la pantalla, por que la persona esté distraída o simplemente por que se haya ido. La idea es que el computador reciba datos de entrada no solo a través de las hábiles manos del usuario que pueda teclear o manejar el ratón prodigiosamente, sino también a través de sus ojos. En este artículo se describen los principales adelantos alcanzados hasta ahora.

DESCRIPCIÓN DEL SISTEMA

Existen dos enfoques que se pueden plantear para realizar el seguimiento de la mirada. El enfoque intrusivo hace referencia a métodos que utilizan dispositivos que deben ser acondicionados al usuario para capturar algunas señales biométricas. En este caso el usuario debe utilizar un tipo especial de gafas o cascos que capturen los movimientos de la cabeza; o mejor, tener implantados electrodos en los ojos; o tal vez, sensores en el cerebro para transmitir las señales eléctricas que se generan, convirtiendo a quien opera el equipo en un usuario biónico.

El método no intrusivo no hace uso de dispositivos físicos sobre o dentro del usuario. Más bien realiza una detección externa capturando imágenes del rostro de la persona, por medio de cámaras de video. El proyecto Mímico sigue este enfoque, utilizando un sistema de visión artificial cuyos componentes básicos son: una fuente de iluminación, un equipo de captura, (que incluye la tarjeta de digitalización y la videocámara) unas rutinas de *software* para procesamiento de imágenes y, por supuesto, el computador (ver figura 1).

Que el computador pueda «ver» al usuario que tiene al frente, determinar hacia dónde está mirando y, entonces, seguir las instrucciones del caso.

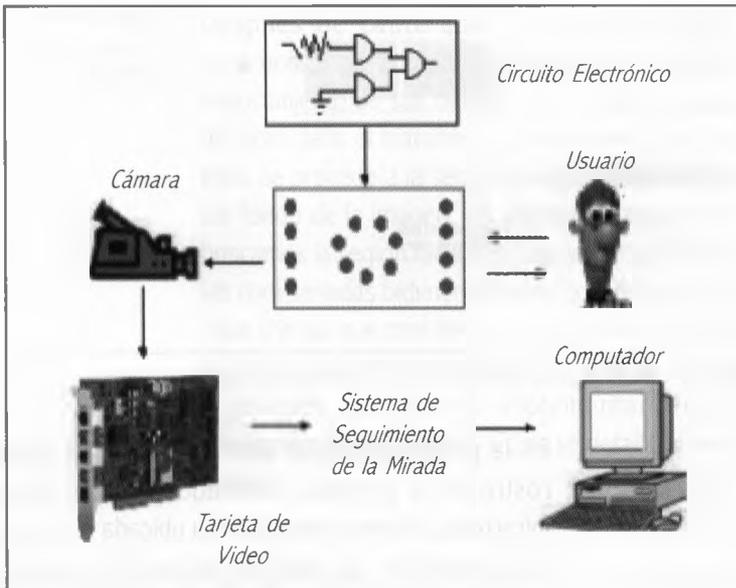


Figura 1. Componentes básicos del sistema.

El modelo típico para sistemas de visión por computador incluye etapas para la captura de video de la escena, preprocesamiento para mejorar los “datos fuente” (imágenes), procedimientos de segmentación para la identificación de los objetos y regiones de interés, y finalmente, extracción de datos para ser usados en la salida computacional^[3]. Fundamentalmente, nuestro proyecto plantea que para realizar el seguimiento de la mirada, se deben tomar imágenes del rostro de la persona y procesarlas para obtener algunas medidas como la ubicación del iris y de la pupila. El modelo propuesto ^[4, 5], consta de las etapas que se observan en la figura 2.

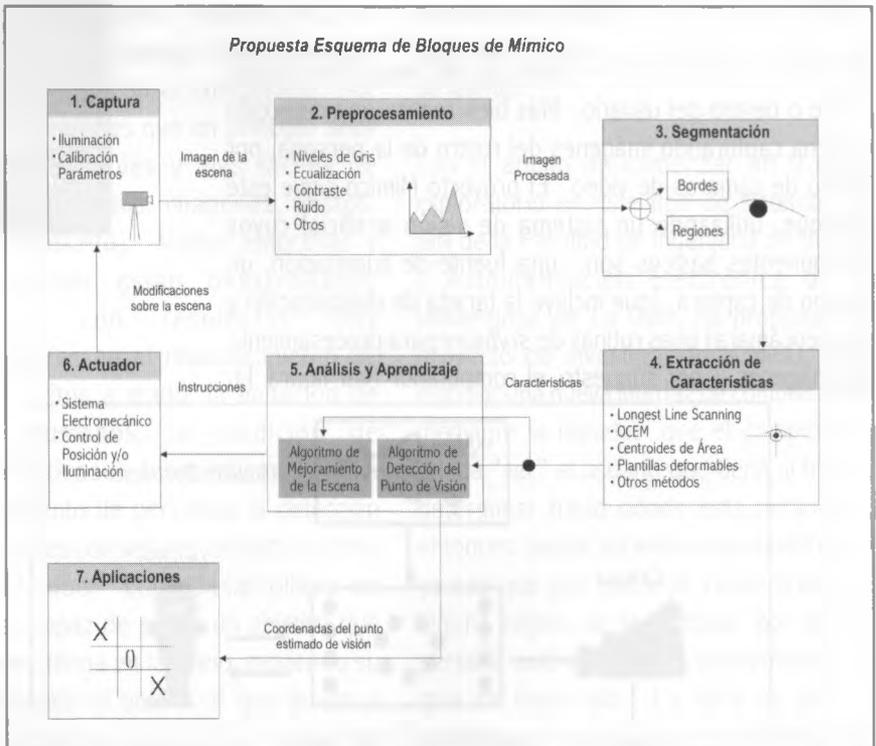


Figura 2. Modelo propuesto para el seguimiento de la mirada.

En la primera etapa se deben capturar las imágenes del rostro de la persona, iluminado con una fuente de luz infrarroja; la cámara se encuentra ubicada cerca a la pantalla del computador. La razón de utilizar infrarrojo es debido a la naturaleza no intrusiva del proyecto, ya que este tipo de

luz no es visible para el ojo humano¹; por esto se puede iluminar sobre él sin que le moleste, y obtener la imagen de pupila brillante, un efecto similar a aquel conocido como el “ojo rojo” que aparece usualmente en las fotografías tomadas con flash (ver figura 3).

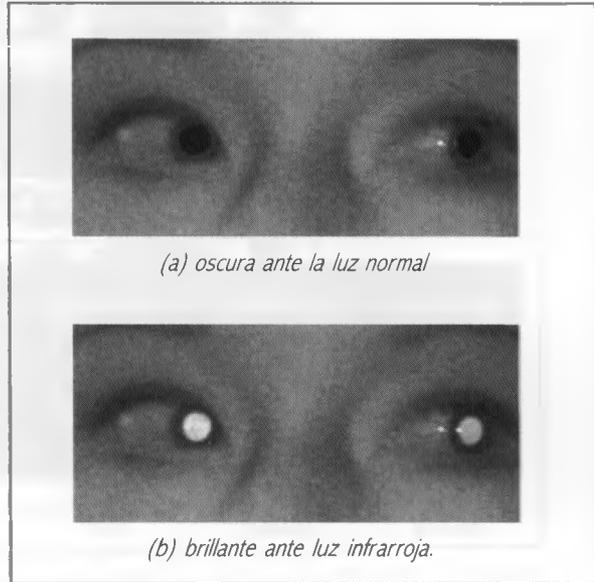


Figura 3. Imágenes de la pupila iluminada.

Después de capturada, la imagen original se debe preprocesar para lograr la eliminación de ruido y hacer un mejoramiento de sus detalles. Para ello se utilizan diversas técnicas para el tratamiento de señales. Una vez realizado esto, se procede a la segmentación, que consiste en separar del fondo de la imagen, los objetos de interés; en este caso buscamos la región facial y ocular, y dentro de ellas se ubican las coordenadas bidimensionales (x,y) del centro de la pupila (que son las que consideramos como una primera estimación de la dirección de la mirada). Con este fin utilizamos varios algoritmos que hemos implementado en diferentes herramientas de *software*. Éstos se detallarán en la siguiente sección.

¹ La longitud de onda de la luz infrarroja se encuentra por encima de los 770nm, límite del espectro de visión humana.

La última etapa del modelo es la que indica que las características extraídas de la mirada (información), puede ser utilizada como entrada para aplicaciones computacionales. Nosotros desarrollamos una aplicación muy simple, que cuenta con varios botones dispuestos en una línea horizontal. El objetivo de la aplicación es permitir que el usuario active alguno de los botones, al mirarlo fijamente. Cada botón indicaría una dirección de la mirada del usuario (si está mirando muy a la derecha, a la derecha, en el centro, a la izquierda, o muy a la izquierda) como se observa en la figura 4.

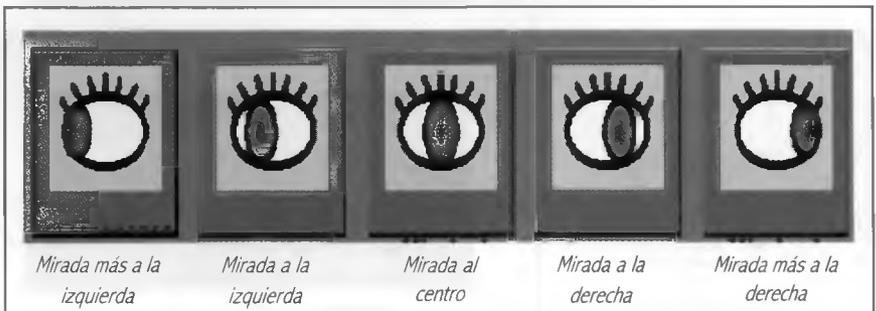


Figura 4. Programa que activa un botón con la mirada de la persona.

RESULTADOS ALCANZADOS

En la actualidad el proyecto se encuentra en la etapa de implementación del modelo explicado anteriormente. Nos hemos concentrado principalmente en el desarrollo de algoritmos y técnicas computacionales para realizar la captura, almacenamiento, preprocesamiento, segmentación, extracción y análisis de las imágenes, utilizando herramientas de *software* como *Java Media Framework* de *Sun Microsystems* ^[6] e *IMAQ Vision* de *National Instruments* ^[7]. Por una parte, hemos obtenido resultados satisfactorios en la detección de rostros y la identificación de regiones faciales y oculares; sobre este punto estamos trabajando con el fin de reducir tiempos de procesamiento y exactitud de las salidas. En la figura 5 se muestra un ejemplo del procesamiento de una imagen.

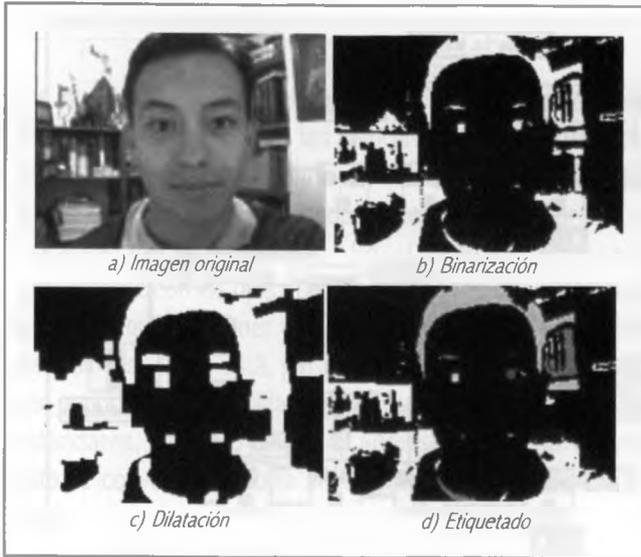


Figura 5. Operaciones de procesamiento de imágenes utilizando JMF.

De otra parte, se trabaja intensivamente en el desarrollo de los algoritmos encargados de analizar las imágenes para detectar la mirada. En este caso se requiere ejecutar filtros, transformaciones y operaciones a nivel de *pixel* y a nivel de imagen. Primero se busca extraer características del ojo, como la ubicación de la pupila, y luego se pretende que el sistema aprenda a clasificar las direcciones en las que posiblemente se orienta la mirada, utilizando técnicas de inteligencia artificial. A manera de ejemplo, en la figura 6 se observa el resultado de uno de los programas que localiza exactamente dentro del ojo la ubicación de la pupila, al comparar la diferencia entre dos imágenes.

Las nuevas técnicas de HCI no intrusivas prometen ser de los principales avances en las ciencias de la computación en este naciente siglo.

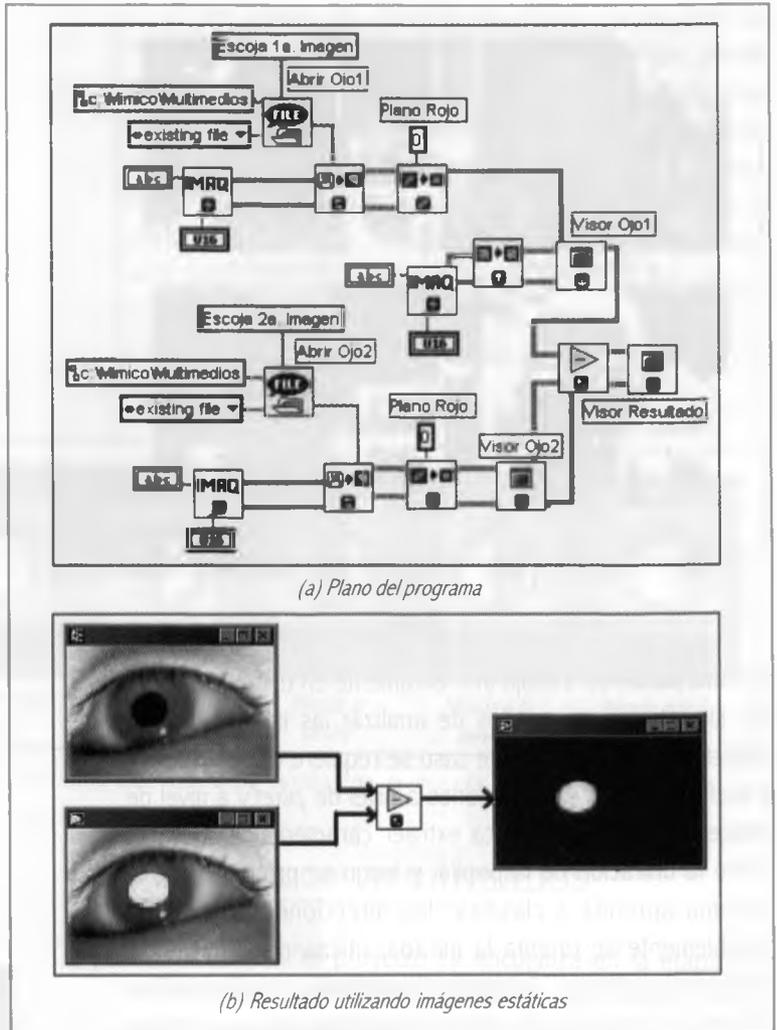


Figura 6. Algoritmo para la detección de la pupila.

Actualmente se está realizando la etapa de experimentación, con el fin de comprobar la confiabilidad del sistema y determinar los ajustes necesarios, ya sea durante la captura de la imagen (se piensa adaptar a la cámara un sistema de rotación y de detección de personas), eliminación de ruido, extracción de características adicionales, o en el aprendizaje del sistema. Los resultados obtenidos hasta el momento han sido expuestos en conferencias dirigidas a los miembros de la Facultad y la Universidad. Se espera realizar su divulgación en algunos simposios internacionales.

CONCLUSIONES

Las nuevas técnicas de HCI no intrusivas prometen ser de los principales avances en las ciencias de la computación en este naciente siglo. Su principal atractivo es que son muy fáciles de usar, ya que se acoplan de manera natural a los usuarios del computador. En el caso de la mirada, simplemente con los ojos una persona podría dar información a un sistema, sin tener que utilizar cascos, gafas u otros dispositivos. Claro está, para que estas interfaces lleguen a ser populares y difundidas, deben basarse en componentes estándares, baratos y de fácil adquisición. Seguramente, esto se convertirá en otra línea de investigación dentro del área.

El hecho de que tengamos las herramientas necesarias para implantar sentidos artificiales al computador, nos hace pensar que los robots futuristas, producto de la ciencia ficción, se conviertan ahora en realidad, y que muy pronto sea cotidiano hablar, escuchar y hasta acariciar a esa maquinita que tenemos sobre el escritorio. Tal vez somos testigos del nacimiento de una nueva era de nuestra civilización informática, en la que el computador tendrá sus ojos bien abiertos. 

BIBLIOGRAFÍA

- [1] [página internet] <http://www.ai.mit.edu/projects/humanoid-robotics-group/cog/cog.html> última visita marzo de 2002.
- [2] [página internet] <http://www.almaden.ibm.com/csl/>, última visita marzo de 2002.
- [3] AWCOCK, G.; THOMAS, R. *Applied Image Processing*, Ed. Mc. Graw Hill, 1996.
- [4] ROJAS, S.A.; RUBIANO, L.; CANTOR, D. "Detección de Rasgos Oculares mediante Visión Artificial", en: *Memorias del II Simposio de Investigación Universidad de La Salle*. Bogotá, D.C., septiembre de 2001.
- [5] ROJAS, S.A., et. al, "Using Human Gaze as Input Data for Computer Programs", en: *International Conference on Computer Science, Software Engineering, Information Technology, e-Business, and Applications (CSITeA'02)*, Iguazú, (Brasil) junio de 2002.
- [6] [página internet] <http://java.sun.com/products/java-media/jmf/2.1.1/guide/>, última visita febrero de 2002.
- [7] National Instruments, *IMAQ Vision User Manual*, 1999.