

## SEGUIMIENTO VISUAL ARTIFICIAL

Entre las capacidades visuales que posee el ser humano se encuentra la de poder seguir objetos que tenemos delante. Aunque lo hagamos de forma inconsciente, mediante esta capacidad podemos realizar múltiples actividades cotidianas, como leer, caminar, conducir, etc. El seguimiento visual está estrechamente imbricado con muchas de las tareas que realizamos. La introducción de estas capacidades en los sistemas visuales artificiales es una de las áreas de investigación en la visión por computador y la robótica de hoy en día. El propósito de este trabajo ha consistido en la introducción de una clase de mejoras que permiten la adaptación automática de los sistemas de visión a entornos variables.

Cayetano Guerra Artal

*Among the several capabilities that human beings own, tracking of objects is one of them. Although we perform it in an unconscious way, by means of this capacity we can carry out multiple daily tasks, like reading, walking, driving, etc. Visual tracking is tightly linked with many of the tasks we do. Nowadays, introducing these capabilities in artificial visual systems is one of the research areas in computer vision and robotics. The aim of this work has been a kind of improvement that allows an automatic adaptation of visual systems to changing environments,*

### PRESENTACIÓN

**E**l seguimiento visual artificial permite a los ordenadores que gobiernan sistemas de visión o robots, disponer de nuevas aplicaciones y capacidades.

### INTRODUCCIÓN

El presente trabajo surgió como respuesta a la necesidad de desarrollar un sistema de detección y reconocimiento de caras mediante técnicas de visión artificial (Aloimonos, 1988). Tempranamente se observó cómo la tarea de identificación de personas requería la captura de buenas vistas frontales de las diferentes caras con las que se experimentaba (Hernández, 1999). La interacción de las personas con el sistema de reconocimiento era completamente natural,

es decir, no se hacía ningún esfuerzo en facilitarle la tarea, con lo que la necesidad de un módulo de seguimiento que mantuviera la cabeza dentro del campo visual del sistema, hasta que el módulo de detección y reconocimiento obtuvieran imágenes válidas que permitieran realizar su tarea, se planteó como parte necesaria en el sistema. No sólo esta aplicación requería de un módulo de seguimiento para su funcionamiento, muchas otras aplicaciones pueden llevarse a cabo o ser mejoradas mediante la utilización de este módulo. Por esto, el trabajo desarrollado plantea como hipótesis la factibilidad de desarrollar módulos de seguimiento autónomos que puedan ser acoplados a sistemas mayores para ofrecer funciones de seguimiento independientemente de su aplicación. Con estas ideas se desarrolló un módulo de seguimiento capaz de actuar e

**Aunque lo hagamos de forma inconsciente, mediante el seguimiento visual podemos realizar múltiples actividades cotidianas, como leer, caminar, conducir, etc.**

**El trabajo desarrollado plantea como hipótesis la factibilidad de desarrollar módulos visuales autónomos.**

integrarse fácilmente con diferentes sistemas y entornos de aplicación.

### **APLICACIONES DEL SEGUIMIENTO**

Existen en la literatura múltiples aplicaciones para el seguimiento visual mediante técnicas de visión artificial. Muchas de ellas están inspiradas en lo que nosotros, como seres humanos, hacemos con nuestro sistema de seguimiento biológico. Sin embargo, a veces resulta difícil darse cuenta de que tareas aparentemente sin nada que ver con el seguimiento, son o están compuestas en alguna medida por seguimiento visual.

El acto de leer es una de ellas, cuando nos fijamos en cada una de las palabras del texto, estamos a la vez siguiendo la línea en donde está incluida, saltando sucesivamente a las palabras situadas a la derecha de la actual.

A continuación se comentan algunas de las aplicaciones donde un módulo de seguimiento tiene aplicación:

- *Navegación*: De la misma manera que las personas utilizan marcas visuales para caminar o evitar objetos, un robot puede hacer uso de ellas para obtener su posición y orientación relativa a ellas durante su trayecto.

Con un sistema de seguimiento multiobjetivo, es posible no sólo usar una simple marca para realizar la tarea, sino varias marcas para obtener una mejora en la fiabilidad y hacer más sencilla la obtención de la posición y orientación.

- *Reconocimiento de caras*: Ésta ha sido una de las aplicaciones implementadas en este trabajo. El reconocimiento de caras es un problema complicado. Existen numerosos trabajos sobre este tema, los cuales tratan de identificar personas

en entornos controlados y con vistas frontales de las caras. El problema se hace más difícil cuando la persona a reconocer no se encuentra inmóvil, sino que se mueve de forma natural delante de las cámaras. DESEO (Hernández, 1999) es una plataforma de visión activa capaz de realizar seguimiento y reconocimiento de personas.

- *Realimentación visual para brazos robot*: Controlar brazos robot por medio de seguimiento visual es una técnica útil para interactuar con entornos dinámicos. Un resumen sobre control visual puede encontrarse en (Hutchinson, 1996).

- *Vigilancia*: Ésta ha sido de las primeras aplicaciones útiles para algoritmos de seguimiento. Normalmente, los sistemas de vigilancia típicos constan de cámaras estáticas con módulos de detección que activan la grabación de secuencias.

- *Seguimiento del movimiento de las personas*: En este tipo de aplicaciones no se sigue a la persona en sí, sino a una parte de su cuerpo. Esto tiene el atractivo de que los gestos del cuerpo y las manos pueden ser reconocidos. Por ejemplo, manos y dedos (Hall, 1999) pueden ser usados como ratones o teclados para interactuar con un ordenador.

- *Control del tráfico*: La monitorización de vehículos en carreteras y autopistas es otra de las aplicaciones prácticas del seguimiento. La detección de colisiones y maniobras peligrosas de los conductores pueden ser detectadas.

### **TÉCNICAS PARA REALIZAR EL SEGUIMIENTO**

El *matching* de patrones es una de las técnicas de visión por computador más antiguas y utilizadas para la identificación o localización de elementos u objetos en imágenes, con una historia de más de

cuatro décadas y una extensa aplicación en diversos campos como: robótica, imágenes médicas, astronomía, proceso de señal de radar, etc.

Su funcionamiento básico es muy simple: un patrón (el cual se define como una imagen, en principio de dimensiones reducidas, que contiene una representación de un objeto o parte de un objeto), es buscado sobre una imagen o parte de una imagen, también llamada ventana de búsqueda.

Para realizar esta búsqueda se desplaza el patrón sobre la imagen, aplicando una función de coincidencia que tomará un valor dependiendo del grado de semejanza entre patrón e imagen.

Las formas de buscar el grado de semejanza entre el patrón y la imagen son múltiples, comenzando por métodos básicos que computan diferencias entre *pixels* hasta esquemas más elaborados, donde en un primer paso los *pixels* son agrupados en primitivas o estructuras más complejas como líneas,

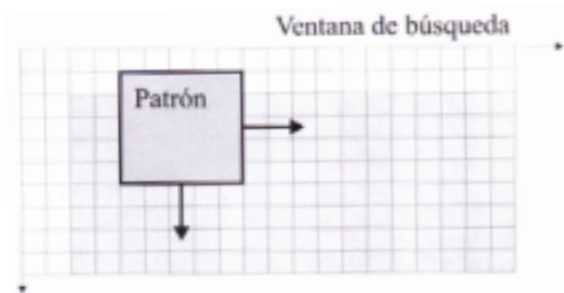


Figura 1: Representación de la imagen y el patrón a buscar dentro de ella.

polígonos, contornos, etc. para posteriormente buscar las coincidencias entre patrón e imagen.

Los algoritmos que computan diferencias entre *pixels* son conceptualmente sencillos, desarrollados a partir de técnicas para el proceso de señal y fundamentados en medidas

estadísticas extensamente estudiadas. Su funcionamiento es simple, como se muestra en la figura 1. Un patrón se desplaza a lo largo de la superficie de la ventana de búsqueda y a medida que recorre todos los *pixels* se genera un valor, gracias a una función de similitud que evalúa el parecido existente entre el patrón y la región de la imagen (del mismo tamaño del patrón) sobre la que se encuentra éste.

Al final del proceso, el punto de máxima similitud será el más parecido y, en principio, será donde se encuentre en la imagen el objeto contenido en el patrón.

### ACTUALIZACIÓN DEL PATRÓN

Es normal que el objeto a seguir varíe su forma o aspecto a lo largo del tiempo.

Este tipo de cambios no sólo corresponden a acercamientos, alejamientos o rotaciones del objeto, sino también pueden ser debidos a deformaciones del mismo, así como cambios en la intensidad luminosa o dirección de la luz.

Todos estos eventos provocan una alteración de los valores de los *pixels* de la imagen que, en teoría, corresponderían a los del objeto de interés.

Puesto que estos valores cambian, el patrón original no sería localizado en los sucesivos *frames*. Para ello, han sido propuestas diversas formas de actualización del patrón original por otro nuevo.

La más utilizada corresponde a la utilización de un umbral que establece un tope que, al ser superado, implica que el patrón ya no se parece lo suficiente al objeto seguido.

El método propuesto en este trabajo comienza con la adquisición de un nuevo patrón tomado de la imagen actual.

La función de similitud de este patrón aplicada sobre la misma imagen devuelve como mínimo un cero, ya que patrón e imagen en este momento son iguales (menor valor implica mayor parecido).

El siguiente paso consiste en obtener el segundo mínimo de la función de similitud; dicho de otra forma, el objeto dentro de la imagen actual más parecido al objeto de interés que se denominará *distractor*. En el tercer paso se calcula el umbral de cambio como la mitad del valor de este segundo mínimo, o lo que es lo mismo, el umbral de cambio se establece a la mitad de distancia entre el primer y el segundo mínimo.

Todo esto se representa en las figuras 2 y 3. En el *frame* a) de la figura 2 se encuentra nuestro objeto de interés marcado con una equis roja en el centro. A su izquierda hay un objeto parecido que generará un segundo mínimo en la función de similitud. Esto se representa en la función a) de la figura 3.

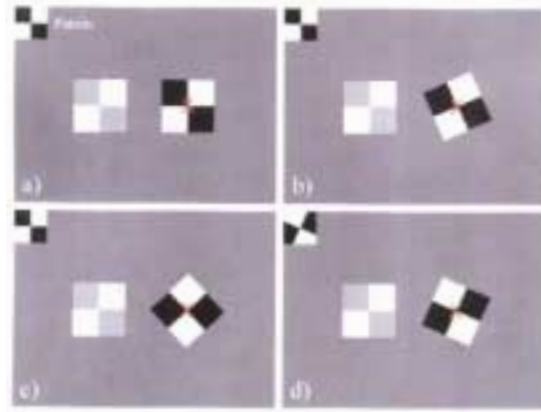


Figura 2: Secuencia donde el objeto a seguir rota y esto hace que su parecido al patrón a buscar sea cada vez menor.

En la esquina superior izquierda de cada *frame* se muestra el patrón que buscamos en cada momento.

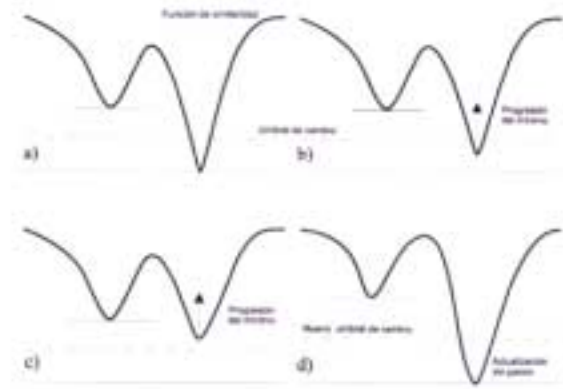


Figura 3: Gráficas de los frames de la secuencia de la figura 2 donde se muestra el valor del mínimo, distractor y umbral de cambio.

En el *frame* b) de la figura 2 se aprecia como el objeto de interés rota y esto provoca que su mínimo ascienda (figura 3, b), puesto que a medida que rota se parece menos al patrón.

Sin embargo, aún no se ha traspasado el umbral de cambio y por tanto no hay actualización.

Durante el *frame* c) se observa que, al haber seguido rotando el objeto, el mínimo traspasa el umbral, con lo que se requiere una actualización del patrón.

En el *frame* d) el patrón ha sido actualizado (el nuevo patrón aparece representado en la esquina superior izquierda del *frame*) y su mínimo es ahora cero (figura 3, d).

De nuevo se busca el segundo mínimo para volver a calcular un nuevo umbral que dependa de lo que rodea al objeto de interés.

La figura 4 muestra un caso real del comportamiento de este método de actualización. La secuencia contiene la base de una puerta que se agranda y gira a medida que la cámara se acerca.

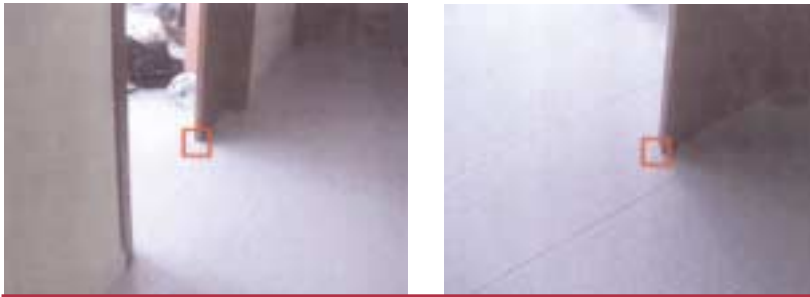


Figura 4: Frames inicial y final de una secuencia de navegación de un robot que utiliza una esquina de una puerta como referencia visual.

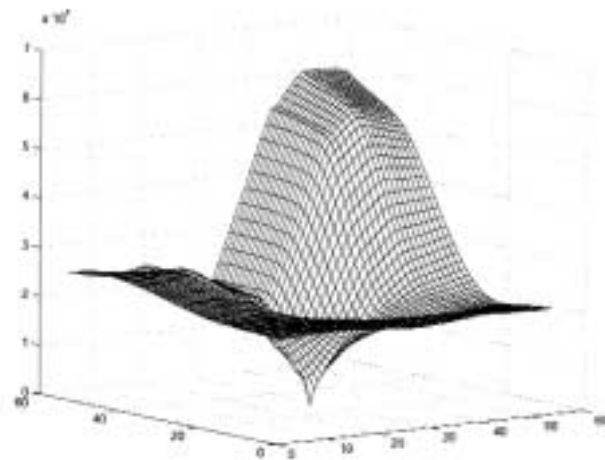


Figura 5: Función de similitud. Su mínimo corresponde con el punto más parecido entre el patrón y la imagen. En otras palabras, donde se encuentra el patrón en la imagen.

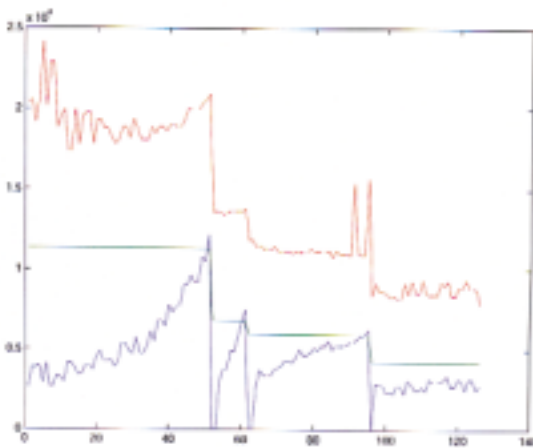


Figura 6: Representación temporal de los valores del mínimo (función azul), el distractor (función roja) y el umbral de cambio de la secuencia.

El objeto de interés en este caso es exactamente la esquina izquierda inferior de la puerta, que aparece marcada con un cuadrado rojo.

Los dos *frames* de la figura corresponden al frame inicial y final de una secuencia de unos cinco segundos.

La figura 5 representa la función de similitud en el primer *frame* de la secuencia. Se puede observar como el mínimo se sitúa en cero.

La figura 6 constituye la progresión del mínimo (función inferior azul), segundo mínimo (función





De esta forma, el desarrollo de módulos autónomos de seguimiento para aplicaciones diversas es posible.

## GLOSARIO

**Matching:** palabra anglosajona que se traduce por concordancia o coincidencia. En este contexto se utiliza para referir el proceso de búsqueda de un patrón sobre la imagen

**OCR:** (Optical Character Recognition) Reconocedor Óptico de Caracteres. Programas que identifican los caracteres escritos en una imagen y los traducen a texto legible por un procesador de textos.

**Pixel:** cada una de las celdillas que definen un punto en una imagen digital.

**Frame:** cada una de las imágenes procedentes de una cámara de vídeo digital. Es el equivalente a un fotograma de una cámara de cine.

## BIOGRAFÍA

### CAYETANO GUERRA ARTAL

Licenciado en informática y doctor en informática por la Universidad de Las Palmas de Gran Canaria.

Profesor asociado en el Departamento de Informática y

Sistemas donde imparte docencia en las asignaturas de

Sistemas Multimedia y Teoría de Autómatas y Lenguajes Formales II.

Edificio Departamental de Informática y Matemáticas  
Campus Universitario de Tafira  
Universidad de Las Palmas de Gran Canaria  
Tlf: 928 458743  
e-mail: cguerra@dis.ulpgc.es

## BIBLIOGRAFÍA

Aloimonos, J. y Weiss, I.: "Active Vision", International Journal of Computer Vision, 1988, págs. 333-356.

Hall, D. y Crowley, J. L.: "Tracking Fingers and Hands with a Rigid Contour Model in an Augmented Reality", 1st International Workshop on Managing Interactions in Smart Environments, 1999.

Hernández, M., Cabrera, J., Castrillón, M., y Guerra, C.: "DESEO: An Active Vision System for Detection, Tracking and Recognition", International Conference on Vision Systems ICVS'99, enero 1999, págs. 376-291.

Hutchinson, S., Hager, G. D. y Corke, P. I.: "A Tutorial on Visual Servo Control", IEEE Trans. Robot. Automat., vol. 12, núm. 5, págs. 651-670.

Patrocinador de esta investigación:

**BELEYMA, S.L.**