

SEGUIMIENTO DE OBJETOS RÍGIDOS CON MOVIMIENTO LIBRE EN SECUENCIAS DE IMÁGENES.

TRACKING FREE-MOVING RIGID OBJECTS IN IMAGE SEQUENCES.

Abril Ayala-Sánchez¹, Patricia Zavaleta-Carrillo, Dámaris Pérez-Cruz.

Fecha de recepción 12 de Enero 2009

Fecha de aceptación 8 de Junio de 2009

RESUMEN

En este trabajo se presenta el desarrollo y la aplicación del proceso para seguir un objeto rígido en movimiento que ha sido grabado previamente en un video (secuencia de imágenes), debido a que el seguimiento no es en tiempo real. El objeto de interés, se encuentra interactuando en un ambiente complejo y en compañía de otros objetos similares o iguales. El proceso de seguimiento se inicia con la localización manual del objeto de interés, por única vez, en el primer cuadro dentro de una escena de la secuencia; continúa con la inicialización de variables del seguimiento; que es la asignación de los primeros valores a las variables del proceso de seguimiento y la definición de las plantillas que son utilizadas posteriormente para predecir el movimiento del objeto y localizar la nueva ubicación que este alcanza. Cuando el proceso de seguimiento se ha iniciado, la nueva localización del objeto se realiza a través de la técnica de plantillas y de diferencia temporal; para finalizar con la predicción del movimiento del objeto y realizar el seguimiento de manera eficiente.

PALABRAS CLAVE: Localización, inicialización, plantillas, diferencia temporal, ambiente complejo.

ABSTRACT

The development and implementation of the process to track a moving rigid object that has been previously recorded in a video (sequence of images) is presented in this work. The object of interest is found interacting in a complex environment and with similar objects. The tracking of the rigid objects starts with the manual location of the object of interest, one time, in the first frame of a sequence; the tracking process follows with the initialization of the variables. This process involves assigning initial values to the tracking variables and defining the templates which are used to predict the motion of the object and find the new location it reaches. When the tracking process has started, the relocation of the object is achieved through the technique of templates and time difference, ending with the prediction and the efficient tracking of the object's motion.

KEYS WORDS: Localization, initialization, templates, time difference, complex environment.

DES-DACH-Área Ciencias de la Información. Universidad Autónoma del Carmen. Calle 56, N° 4, esquina Av. Concordia. Col. Benito Juárez. C.P. 24180
Ciudad del Carmen, Campeche, México.

Autor para correspondencia: aayala@pampano.unacar.mx.

INTRODUCCIÓN

El seguimiento de un objeto dentro de una secuencia de imágenes es un reto al que se enfrenta el área de Inteligencia Artificial, particularmente el tratamiento de imágenes, ya que esto representa un problema a resolver. El seguir un objeto específico en una secuencia de imágenes donde existen otros objetos que también forman parte de la escena complica la acción. Esto ha dado lugar a desarrollar alternativas de solución que permitan alcanzar un mayor nivel de eficiencia en el seguimiento.

El desarrollo del trabajo que se presenta corresponde a una propuesta de solución eficiente para seguir un objeto rígido que está en un ambiente complejo, -fondo de la escena donde se encuentra interactuando el objeto-, compuesto por otros objetos, incluso, semejantes a él.

Para alcanzar una solución eficiente al problema que se plantea en este trabajo, se utilizan de forma conjunta técnicas que permiten alcanzar la segmentación del objeto de interés y procedimientos propios de la investigación. De tal forma es conveniente recordar que los métodos basados en características siguen facetas individuales como puntos (Zheng and Chellappa, 1995; Beymer et al., 1997), en líneas (Deriche and Faugeras, 1990; Liu and Huand, 1991), o curvas (Zhang and Faugeras, 1992) normalmente basándose en esquemas de correspondencia (Badenas et al., 2000), lo cual no es el caso de esta aplicación, ya que estamos considerando la correspondencia entre regiones del objeto de interés basándonos en plantillas (Coto, 2003; Myler and Weeks 1993).

El proceso de seguimiento considera métodos de mejora de la imagen para poder obtener resultados más precisos para el se-

guimiento, esto se logra aplicando el filtro pasabajo (González and Woods, 1992), que mejora la imagen y aumenta la eficiencia al proceso de segmentación.

Dado que, según las técnicas de seguimiento pueden ser agrupadas en cuatro categorías: métodos basados en información tridimensional, métodos basados en características, métodos basados en modelos deformables y métodos basados en regiones (Badenas et al., 2000), aquí se considera la región definida por el usuario, que es la que proporciona los datos necesarios para realizar la inicialización de las variables del seguimiento.

Realizar una segmentación lo más precisa posible es determinante para lograr que el seguimiento del objeto se realice de forma eficiente. Es importante mencionar que para el problema de oclusión, que se presenta cuando el objeto seguido se encuentra con otros objetos que están delante de él (lo cubren parcial o totalmente), no se han considerado soluciones; el sistema de visión por computadora y la ubicación del lugar donde se lleva a cabo el seguimiento de objetos, puede representarse de manera gráfica (Figura 1)

MATERIALES Y METODOS

IDENTIFICACIÓN MANUAL DEL OBJETO

El proceso de solución al problema de seguir un objeto se inicia con la selección del objeto que se desea seguir. Esto se hace manualmente a través del ratón (figura 2)

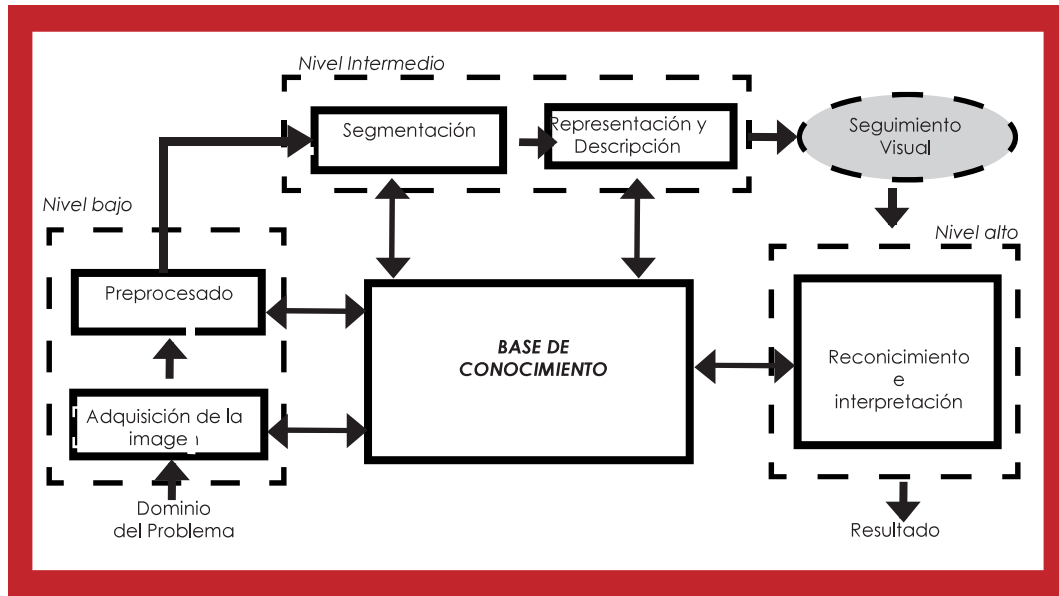


Figura 1. Sistema de Visión por Computadora que muestra la ubicación del proceso de seguimiento (en color gris). (González 1996)

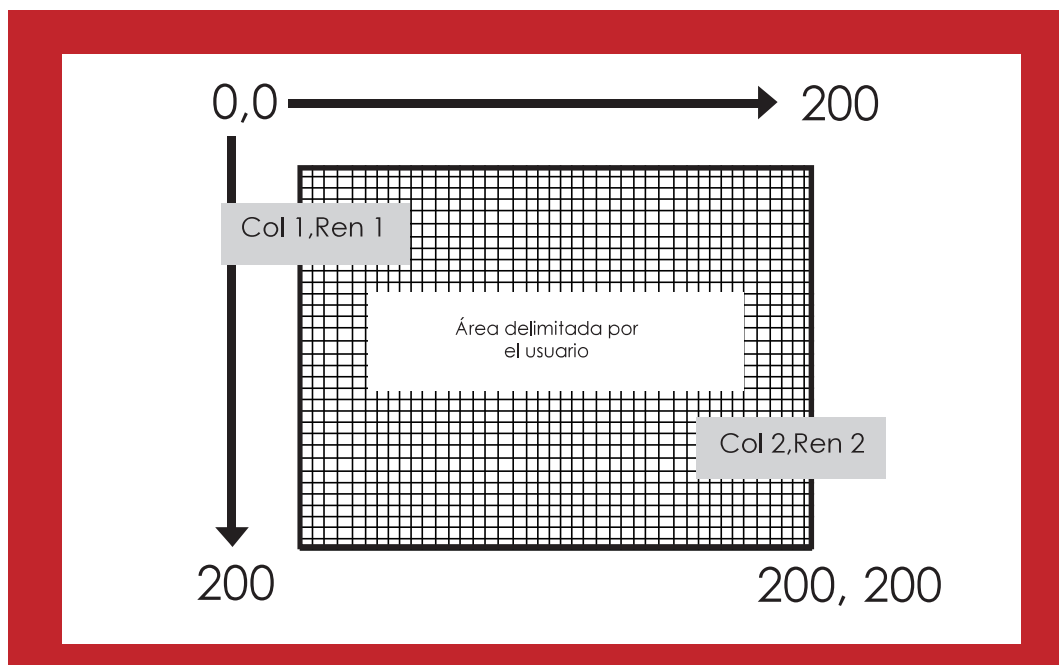


Figura 2. Área delimitada por el usuario que contiene el objeto de interés.

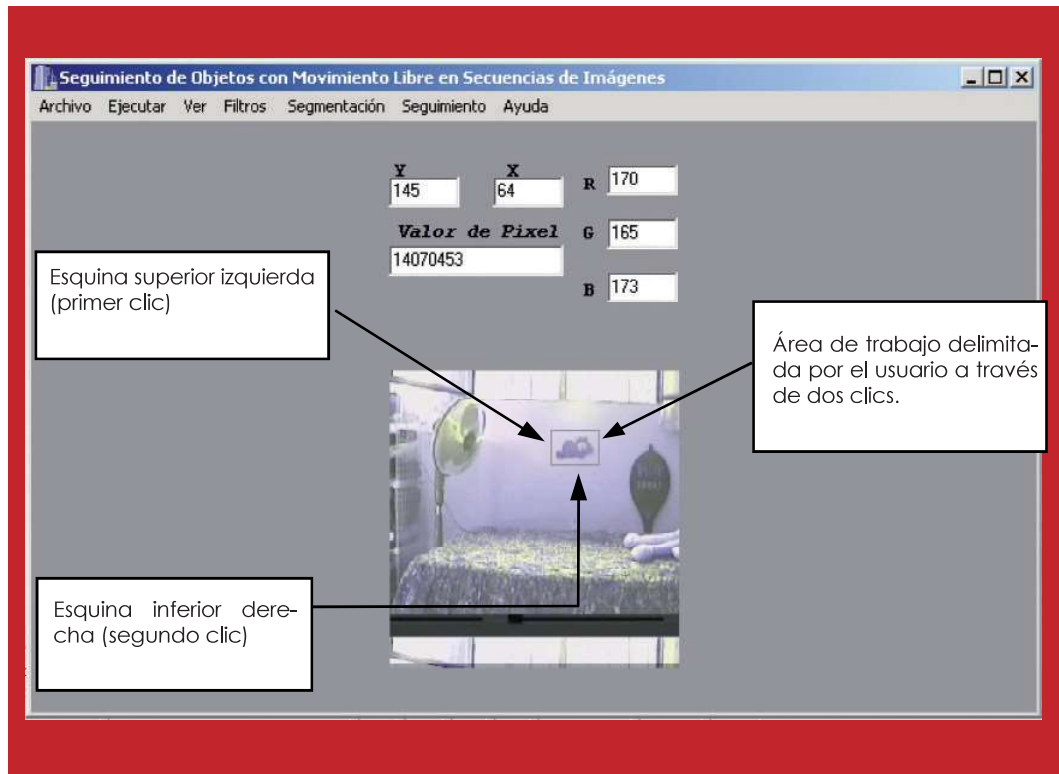


Figura 3. Delimitación del área de trabajo y localización manual del objeto.

El área de interés delimitada por el usuario corresponde a solo una parte del primer cuadro de la secuencia de imágenes (figura 3), que muestra dentro de un cuadro de una escena compleja el área que limita al objeto que se va a seguir.

Cuando se ha definido el perímetro, que abarca en su totalidad al objeto, se realiza el proceso de segmentación para separar el objeto de interés de los demás objetos existentes dentro del área. En esta parte del proceso de seguimiento se identifica la región de importancia donde se encuentra el objeto de interés y se lleva a cabo la segmentación, a través del etiquetado de regiones, para aislar al objeto de interés,

considerando de antemano la aplicación del filtro pasa bajo para suavizar la imagen y resaltar al objeto de interés y así poder obtener una segmentación más precisa (figura 4).

La segmentación de una imagen o porción de imagen, como es en este caso, se puede interpretar como la separación de esa imagen en regiones (áreas) no traslapadas y homogéneas entre sí, con respecto a alguna característica. Si decimos que el dominio de la imagen está dado por I , entonces el problema de segmentación consiste en determinar el conjunto $S_k \subset I$ cuya unión es la imagen I completa. Por lo tanto, el conjunto que conforma la segmentación debe satisfacer:

$$(1) \quad I = \bigcup_{k=1}^K S_k$$

Donde $S_k \cap S_j = \emptyset$ para $k \neq j$, y cada S_k está conectado. Lo ideal es que un método de segmentación encuentre aquellos conjuntos que corresponden a distintas estructuras o regiones anatómicas de interés en la imagen.

Al eliminar la restricción de que las regiones estén conectadas, determinar los conjuntos S_k es llamado **clasificación de píxel** y a los conjuntos se les llama **clases**. La clasificación de píxeles frecuentemente es un

objetivo deseable en el tratamiento de imágenes. La determinación del número de clases K en la clasificación de píxeles puede ser un problema complejo, por lo que generalmente se asume conocida.

El **etiquetado** es el proceso de asignar una identificación significativa (como el color, área, forma, entre otras) a cada región y puede ser llevada a cabo separadamente de la segmentación. Este proceso mapea el índice numérico del conjunto S_k , a una designación anatómica.

INICIALIZACIÓN DEL SEGUIMIENTO

Cuando ya está definida la primera plantilla que se utiliza para realizar el seguimiento y que representa los valores iniciales del objeto, se utiliza como base para obtener la segunda plantilla que surge al hacer un co-



Figura 4. Aplicación del filtro pasa bajo al área de interés.

rimiento de los cuadros de la secuencia de imágenes, para este caso un salto de tres cuadros calculados de forma empírica al probar con secuencias que contienen objetos rígidos que interactúan con otros objetos similares. Se aplica el mismo proceso definido anteriormente; a la misma región seleccionada se le aplica un proceso de segmentación y se separa al objeto para obtener la segunda plantilla, esto nos sirve para comparar y comprobar que la primera plantilla definida coincide.

Los métodos basados en plantillas (*atlas-guided methods*) son una poderosa herramienta para la segmentación de imágenes (generalmente utilizadas en aplicaciones médicas), cuando esta disponible una plantilla o mapa estándar como también se les suele llamar. La plantilla se genera por información obtenida de la imagen que requiere segmentarse. La plantilla se utiliza como un marco de referencia para segmentar nuevas imágenes, en cuadros o imágenes subsecuentes de la secuencia de imágenes. En una interpretación conceptual; los métodos basados en plantillas son similares a los clasificadores, con la excepción de que están implementados en el dominio espacial de la imagen en lugar de en un espacio característico. Los métodos basados en plantillas primero encuentran una transformación uno a uno que transforma la plantilla de la imagen pre-segmentada a la imagen deseada que requiere segmentación, a este proceso se le conoce como una deformación de plantilla (*atlas warping*). La deformación puede ser realizada usando transformaciones lineales, pero debido a la variabilidad anatómica, frecuentemente se utiliza una aplicación secuencial de transformaciones lineales y no lineales.

Después de haber obtenido las plantillas del objeto, el paso siguiente es identificar la ubicación del objeto. Esto se realiza a través de la localización de los cuatro puntos máximos extremos del objeto que se cono-

cen como coordenadas; la coordenada inferior, coordenada superior, coordenada izquierda y coordenada derecha. Obtener estas coordenadas representa para el seguimiento una parte muy importante a la cual llamamos inicialización de variables del seguimiento porque, una vez iniciado el proceso, el movimiento del objeto se determina a partir del aumento o disminución de cada una de las coordenadas que fueron identificadas. Esto permite distinguir el contorno del objeto de interés con un color diferente al del objeto y de esta forma poder visualizar como el objeto va moviéndose en la secuencia de imágenes

LOCALIZACIÓN DEL OBJETO

Cuando las variables obtienen su primer valor (valor inicial), y se han definido las plantillas, el proceso de inicialización está completo y se puede llevar a cabo el proceso de localización del objeto, que consiste en predecir su movimiento. Esto se consigue a través de calcular la diferencial temporal a través de los datos obtenidos en el proceso de segmentación (coordenadas) y hacer coincidir las plantillas, al hacer la correspondencia entre regiones, a través del método de plantillas, para localizar al objeto que se esta siguiendo. De esta forma se estima el movimiento que tuvo el objeto de interés.

TÉCNICAS DE PREDICCIÓN EN LA LOCALIZACIÓN DE OBJETOS

La predicción, en este caso, significa estimar hacia donde se va a mover un objeto; en el seguimiento podría no ser tan complicado cuando se habla de objetos que tienen un movimiento lineal, esto es conocer hacia donde se dirige y no existe entonces, la necesidad de predecir. Cuando el objeto que se sigue tiene movimiento libre, el seguimiento se complica, ya que esto significa predecir cuál será la dirección hacia donde se moverá el objeto, o mejor dicho cuales serán las posibles direcciones hacia las cuales se puede mover el objeto de interés.

La predicción se lleva a cabo aplicando una correlación discreta entre plantillas. La cual consiste en el proceso de comparar matemáticamente una imagen con otra, para dar como resultado una expresión en bidimensional de equivalencias (matricial). La ecuación que define la correlación discreta bidimensional se representa:

(2)

$$Out(i,j)=\sum_{m=1}^M \sum_{n=1}^N In1((m,n))In2(i+m, j+n)$$

Donde **In1** e **In2** son las imágenes de entrada y **Out** es la imagen de salida. La dimensión de la imagen es M por N, que recordemos que lo define el usuario para esta aplicación, y las imágenes de entrada pueden ser de tamaños diferentes ya que la imagen más pequeña será completada con ceros para poder realizar la convolución de manera consistente.

El procedimiento de relocalización del objeto se realiza aplicando diferencia temporal entre plantillas, utilizando como valores de inicio las coordenadas obtenidas en la parte de inicialización y las plantillas definidas. Cuando se ha cambiado de cuadro, dentro de la secuencia de imágenes, se estima el movimiento que ha ocurrido de un cuadro a otro aplicando la diferencia temporal entre plantillas y se cambian los valores en los parámetros de las coordenadas. Una vez realizada la actualización de los valores de las coordenadas se calcula el movimiento que posiblemente existirá en el siguiente cuadro a evaluar.

Los movimientos que realiza el objeto que se sigue regularmente están dados en una proporción que varía en un rango de dos a cinco píxeles. Cuando el movimiento se

sale de este rango, definido en la parte de inicialización del seguimiento, debe realizarse un ajuste de valores aplicando nuevamente la inicialización del seguimiento de la misma forma que se realiza en la parte de inicialización del seguimiento. Se actualizan los valores de las coordenadas y los valores de las plantillas y se continúa el seguimiento a través del algoritmo descrito.

El movimiento del objeto en cada cuadro procesado de la secuencia de imágenes se puede apreciar de manera visual a través de marcar el contorno de la región en cada cuadro de la reproducción del video, siguiendo al objeto (figura 4). Una vez que se ha diferenciado el contorno del objeto de interés con un color negro, se realiza un corrimiento en los cuadros de la secuencia de imágenes para obtener una nueva imagen y aplicar los distintos procesos para estimar los nuevos valores de movimiento y conseguir de esta forma hacer el seguimiento del objeto.

Es importante mencionar que la imagen procesada corresponde a la sección delimitada por el usuario al inicio del seguimiento. Esta región puede ser una parte o la totalidad del área dispuesta para mostrar la secuencia de imágenes.

Los métodos basados en correspondencia y plantillas son solamente algunos de los muchos utilizados, según el estado del arte, y que han mostrado resultados eficientes en distintas aplicaciones de seguimiento de objetos.

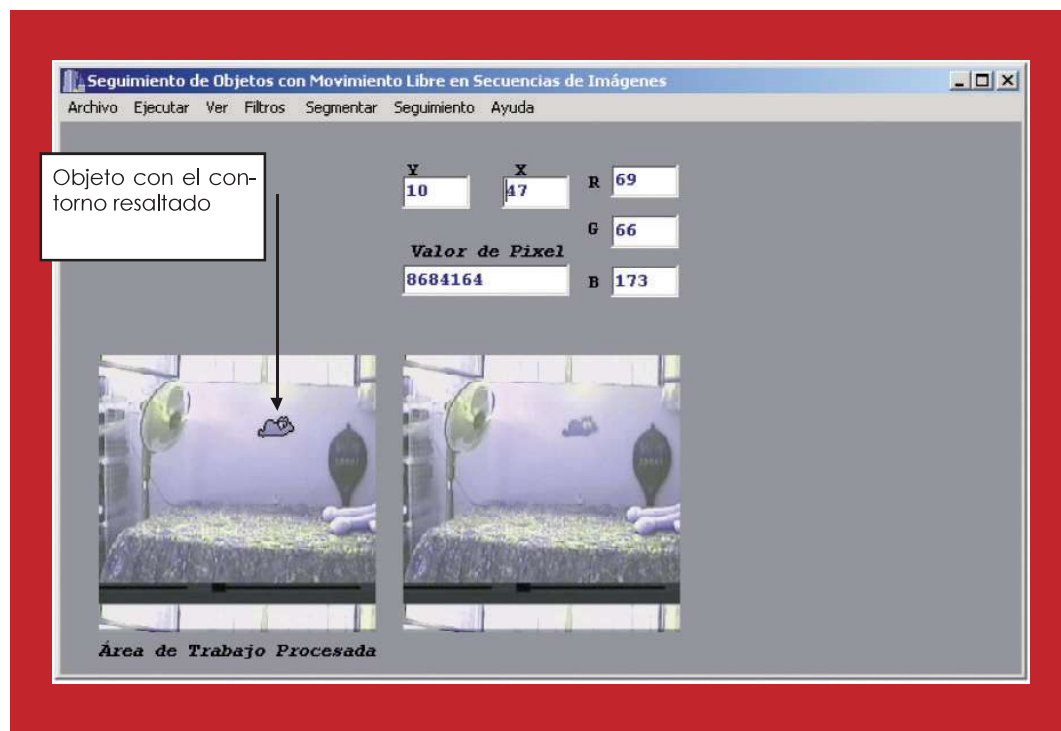


Figura 5. El cuadro izquierdo muestra el contorno del objeto resaltado para hacer notar que es al que se le va a dar seguimiento. Fuente: Elaboración propia

RESULTADOS

Se han considerado diversas referencias que permitieron conocer técnicas utilizadas para resolver el problema de seguimiento. A su vez, tener un panorama más amplio para poder plantear un algoritmo de solución al problema de seguimiento de objetos rígidos.

Cuando se aplica el algoritmo de solución propuesto en escenas que involucran objetos alfaflexibles los resultados obtenidos son similares a seguir un objeto rígido cuando

no sufre cambios bruscos ya que al momento de ir cambiando su forma, considerando que lo hace lentamente, puede realizarse una actualización de la información que modifique el contenido de la información de las plantillas y, por consecuencia, identificar y continuar el seguimiento del objeto que se ha seguido de inicio.

Para las consideraciones tomadas al realizar la segmentación del objeto, como aplicar el filtro y determinar el umbral para poder realizar una binarización del área de interés, se puede comprobar que cuando esto se realiza de forma precisa y la segmentación

del objeto es clara, el seguimiento se lleva a cabo de forma regular, sin mostrar alteraciones o perder al objeto. Recordando que en este caso se están considerando escenas en las cuales la complejidad no es excesiva, no existe oclusión u otros objetos que interfieran con el objeto de interés, por tanto, la segmentación se obtiene de forma clara y conveniente para realizar un seguimiento satisfactorio.

Se puede observar claramente que en la etapa de segmentación, es conveniente utilizar o desarrollar técnicas que permitan obtener mejores resultados en escenas muy complejas, básicamente aquellas que contienen oclusión del objeto de interés.

Dentro del total de pruebas definidas para la herramienta se utilizaron diversos objetos rígidos, tales como esferas en distintos tamaños, frutas y entre ellos se considero una mano, que es un objeto alfaflexible. Se pudo apreciar que la segmentación del objeto se hace muy lenta y el algoritmo tiende a actualizar la plantilla constantemente, ya que como se trata de un objeto deformable que va cambiando constantemente, esto lo detecta el sistema y actualiza de forma muy constante las plantillas. De esta forma podemos concluir que aún cuando realiza el seguimiento de un objeto alfaflexible, no es el mejor algoritmo de solución, ya que el tiempo empleado en el proceso es demasiado.

De las técnicas de inicialización del seguimiento se utiliza el de aprendizaje de la escena que permite obtener del objeto de interés en la escena, las características más relevantes y descriptivas que funcionan para realizar de manera eficiente la diferencia temporal entre plantillas. El éxito del proceso de seguimiento está condicionado

a la exactitud con que se hayan segmentado las regiones de los cuadros y por consiguiente la exactitud de las plantillas obtenidas.

La comparación del objeto en los cuadros analizados resulta satisfactoria ya que las plantillas son actualizadas dado un número determinado de cuadros del video.

La diferencia temporal y las plantillas, al utilizarse a la par, reducen el trabajo del algoritmo al calcular una diferencia temporal solo en puntos estratégicos de la plantilla (las coordenadas), donde las plantillas representan al objeto de interés.

Para efectos de esta propuesta de solución, las pruebas que se realizaron para corroborar la etapa de segmentación demuestran que cuando la escena no es compleja, los resultados son excelentes para aislar al objeto de interés de los demás objetos existentes. Pero cuando se trabajó con escenas muy complejas, la segmentación se concluía poco clara y como no se podía aislar de forma clara al objeto de interés, el seguimiento se perdía.

CONCLUSIONES

Basado en las pruebas que se realizaron utilizando fondos complejos, objetos similares al de interés, podemos comprobar que el algoritmo propuesto funciona eficientemente al seguir objetos rígidos que se encuentran en una secuencia de imágenes (video).

Cuando existe un ambiente complejo, objetos similares, no existe oclusión y el objeto de interés se segmenta totalmente den-

tro de la escena, el seguimiento se lleva a cabo de forma correcta, sin perder al objeto de interés, durante toda la secuencia de imágenes.

Cuando la segmentación es deficiente, el seguimiento puede empezar bien pero es incierto el rumbo que tomará, ya que el algoritmo tiende a actualizar la información de las plantillas considerando cierto porcentaje de diferencia y si ocurre muy frecuentemente pueden ocurrir dos cosas: que ajuste la segmentación al objeto de interés o que cambia totalmente la plantilla del objeto.

En las secuencias de imágenes donde se seguía un objeto que en algún momento de la secuencia se traslapaba, el seguimiento se veía interrumpido, como está considerado, ya que el algoritmo de solución no considera la oclusión y al encontrar una diferencia entre las plantillas más altas a la considerada permitida, simplemente considera perdido al objeto de interés.

BIBLIOGRAFÍA

- Badenas, J.; Sanchiz, J.M. and Pla, F. 2000. Using temporal integration for tracking regions in traffic monitoring sequences. *Proceedings of the International Conference on Pattern Recognition*. 3: 1125-1128.
- Beymer D.; McLauchlan P.; Coifman B. and Malik J. 1997. A real-time computer vision system for measuring traffic parameters. *Proceedings., IEEE Computer Society Conference*. 17:495 – 501
- Coto, E. 2003. *Métodos de segmentación de imágenes médicas, Lecturas en ciencias de la computación*. Universidad Central de Venezuela. Facultad de Ciencias. Escuela de Computación. Laboratorio de Computación Gráfica. Venezuela, Caracas. ISSN 1316-6239.
- Deriche R.; and Faugeras O. 1990. Tracking line segments. *Image and Vision Computing*, 8(4):261-270.
- González, R. C.; and Woods, R. D. 1996. *Tratamiento Digital de Imágenes*. Addison-wesley/Díaz de Santos. 800 p.
- Liu Y. and Huand T.S. 1991. Determining straight line correspondences from instensity images. *Pattern Recognition*, 24(6):489-504.
- Myler, H. R. and Weeks, A. R. 1993. *The pocket handbook of image processing algorithms in C*. Prentice Hall. 303 p.
- Zhang Z. and Faugeras O.D. 1992. Theree-Dimensional motion computation and object segmentation in a long sequence of stereo frames. *Int. J. Comp Vision*. 7(3):211-241.
- Zheng Q. and Chellappa. 1995. R. Automatic feature point extraction and tracking in image sequences for arbitrary camera motion. *Int. J. Comp Vision*. 15(2):31-76.

Forma correcta de citar este trabajo:

Ayala-Sánchez, A., Zavaleta-Carrillo, P. y Pérez-Cruz, D. 2009 Seguimiento de objetos rígidos con movimiento libre en secuencias de imágenes. *U. Tecnociencia* 3 (1) 24 - 34.