

SEGUIMIENTO DE OBJETOS MÓVILES USANDO UNA CÁMARA PTZ

T.G. 0984

CAROLINA PINZÓN FLÓREZ

**DIRECTOR:
ING. FRANCISCO CARLOS CALDERÓN BOCANEGRA**

**ASESOR:
ING. ALEJANDRO FORERO GUZMÁN**

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE ELECTRÓNICA
BOGOTÁ D.C, JUNIO 2010**

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA ELECTRÓNICA**

RECTOR MAGNÍFICO: PADRE JOAQUÍN EMILIO SÁNCHEZ GARCÍA S.J

DECANO ACADÉMICO: ING. FRANCISCO JAVIER REBOLLEDO MUÑOZ

DECANO DEL MEDIO UNIVERSITARIO: PADRE SERGIO BERNAL RESTREPO S.J

DIRECTOR DE CARRERA: ING. JUAN MANUEL CRUZ BOHÓRQUEZ

DIRECTOR DE PROYECTO: ING. FRANCISCO CARLOS CALDERÓN BOCANEGRA

ASESOR DEL PROYECTO: ING. ALEJANDRO FORERO GUZMÁN.

ARTICULO 23 DE LA RESOLUCIÓN No. 13 DE JUNIO DE 1946

“La Universidad no se hace responsable de los conceptos emitidos por sus alumnos en sus proyectos de grado.

Solo velará porque no se publique nada contrario al dogma y a la moral católica y por que los trabajos no contengan ataques o polémicas puramente personales. Antes bien, que se vea en ellos el anhelo de buscar la verdad y la justicia.”

TABLA DE CONTENIDO

ÍNDICE DE FIGURAS	6
ÍNDICE DE TABLAS.....	8
1. INTRODUCCIÓN.....	9
2. MARCO TEÓRICO.....	11
2.1 FORMACIÓN Y REPRESENTACIÓN DE UNA IMAGEN.....	11
2.2 RECONOCIMIENTO DE PATRONES EN UNA IMAGEN.....	11
2.2.1 CLASIFICACIÓN DE DISTANCIA MÍNIMA.....	12
2.2.2 ALGORITMOS DE COMPARACIÓN “ <i>MATCHING TEMPLATE</i> ”.....	13
2.2.2.1 MÉTODO DE DIFERENCIA CUADRADA.....	13
2.2.2.2 MÉTODO DE CORRELACIÓN.....	14
2.2.2.3 MÉTODO DE COEFICIENTE DE CORRELACIÓN.....	14
2.3 SEGUIMIENTO DE OBJETOS EN MOVIMIENTO	15
2.3.1 SEGUIMIENTO DE OBJETOS BASADO EN FILTRO DE PARTÍCULAS.....	16
2.4 CÁMARA SONY EVI-D70	17
2.4.1 ZOOM.....	18
2.4.2 LISTA DE COMANDOS	18
2.4.3 PAN - TILT.....	19
3. ESPECIFICACIONES	20
3.1 DIAGRAMA GENERAL.....	20
3.2 ESPECIFICACIONES GENERALES.....	20
3.3 DESCRIPCIÓN.....	21
3.4 DIAGRAMA DE FLUJO DEL MÉTODO DE FILTRO DE PARTÍCULAS	22
3.5 DIAGRAMA DE FLUJO DEL MOVIMIENTO DE LA CÁMARA	25
4. DESARROLLOS	26

4.1	DESARROLLO DEL ALGORITMO DE FILTRO DE PARTÍCULAS	26
4.2	DESARROLLO DEL ALGORITMO DEL PILOTO DE LA CÁMARA PTZ	29
4.3	PROTOCOLO DE PRUEBAS	31
5.	ANÁLISIS DE RESULTADOS.....	33
6.	CONCLUSIONES.....	55
7.	BIBLIOGRAFÍA	57
8.	ANEXOS	59
	ANEXO A COLOR	59
	ANEXO A.1 MODELO DEL COLOR RGB.....	59
	ANEXO A.2 MODELO CMYK.....	60
	ANEXO A.3 MODELO HSV	61
	ANEXO B SEGMENTACIÓN DE UNA IMÁGEN	61
	ANEXO C RECONOCIMIENTO DE PATRONES EN UNA IMAGEN.....	62
	ANEXO D ALGORITMO DE FILTRO DE PARTÍCULAS (MÉTODO DE COEFICIENTE DE CORRELACIÓN)	63
	ANEXO E ALGORITMO DE FILTRO DE PARTÍCULAS (MÉTODO DE DIFERENCIA CUADRADA)	63
	ANEXO F ALGORITMO DE PILOTO DE LA CÁMARA PTZ.....	63
	ANEXO G ALGORITMO PRUEBAS – PUNTOS A MANO.....	63
	ANEXO H ALGORITMO PRUEBAS – COEFICIENTE DE CORRELACIÓN	63
	ANEXO I ALGORITMO PRUEBAS – DIFERENCIA DE CUADRADOS.....	63
	ANEXO J RESULTADOS PRUEBAS DE SEGUIMIENTO	63

ÍNDICE DE FIGURAS

Figura 1 Diagrama de bloques de la decisión teórica de la clasificación de patrones [6]	11
Figura 2 Actualización y Predicción	16
Figura 3 CÁMARA SONY EVI-D70 [4].....	17
Figura 4 Diagrama de Bloques general [16],[4].....	20
Figura 5 Diagrama de flujo del Filtro de Partículas	24
Figura 6 Diagrama de flujo del movimiento de la cámara	25
Figura 7 Selección Manual de la Región de Interés	26
Figura 8 Región de Interés ROI seleccionada	27
Figura 9 Ventana con un largo y un ancho determinado.....	27
Figura 10 Generación de los rectángulos aleatorios en el cuadro $k = n+1$	28
Figura 11 Coordenadas de una imagen.....	30
Figura 12 Distancia en x y en y entre el punto central de la región y el punto central del cuadro	30
Figura 13 Grafica comparación puntos Manuales y Automáticos	33
Figura 14 Distancia Euclidiana entre puntos Manuales y Automáticos.....	34
Figura 15 Grafica comparación puntos Manuales y Automáticos	35
Figura 16 Distancia Euclidiana entre puntos Manuales y Automáticos.....	36
Figura 17 Grafica comparación puntos Manuales y Automáticos	37
Figura 18 Distancia Euclidiana entre puntos Manuales y Automáticos.....	37
Figura 19 Grafica comparación puntos Manuales y Automáticos	38
Figura 20 Distancia Euclidiana entre puntos Manuales y Automáticos.....	38
Figura 21 Grafica comparación puntos Manuales y Automáticos	39
Figura 22 Distancia Euclidiana entre puntos Manuales y Automáticos.....	40
Figura 23 Grafica comparación puntos Manuales y Automáticos	41
Figura 24 Distancia Euclidiana entre puntos Manuales y Automáticos.....	41
Figura 25 Grafica comparación puntos Manuales y Automáticos	42
Figura 26 Distancia Euclidiana entre puntos Manuales y Automáticos.....	42
Figura 27 Grafica comparación puntos Manuales y Automáticos	43
Figura 28 Distancia Euclidiana entre puntos Manuales y Automáticos.....	44
Figura 29 Grafica comparación puntos Manuales y Automáticos	45
Figura 30 Distancia Euclidiana entre puntos Manuales y Automáticos.....	45
Figura 31 Grafica comparación puntos Manuales y Automáticos	46

Figura 32 Distancia Euclidiana entre puntos Manuales y Automáticos.....	46
Figura 33 Grafica comparación puntos Manuales y Automáticos	47
Figura 34 Distancia Euclidiana entre puntos Manuales y Automáticos.....	48
Figura 35 Grafica comparación puntos Manuales y Automáticos	49
Figura 36 Distancia Euclidiana entre puntos Manuales y Automáticos.....	49
Figura 37 Grafica comparación puntos Manuales y Automáticos	50
Figura 38 Distancia Euclidiana entre puntos Manuales y Automáticos.....	50
Figura 39 Grafica comparación puntos Manuales y Automáticos	51
Figura 40 Distancia Euclidiana entre puntos Manuales y Automáticos.....	52
Figura 41 Representación del modelo RGB [11], [12].....	59
Figura 42 Modelos RGB y CMYK [13].....	60
Figura 43 a) Tono del color (Hue) (b) Saturación (S) (c) Luminancia (d) Modelo HSV	61

ÍNDICE DE TABLAS

Tabla 1. Características principales de la cámara EVI_D70 [4].....	18
Tabla 2 parámetros de RS-232 y RS-422 [4]	19
Tabla 3 Movimientos de PAN y TILT [4].....	19
Tabla 4 Posibles objetos de interés a seguir	31
Tabla 5 Fondos	32
Tabla 6 Tabla de Tiempos	54

1. INTRODUCCIÓN

En la actualidad se han desarrollado diferentes sistemas de automatización los cuales realizan algunas de las labores que ejercen las personas. Entre todos estos sistemas se encuentra el seguimiento automático de objetos móviles, el cual es un tema de gran interés ya que unas de sus aplicaciones más importantes y conocidas son la seguridad, los deportes [1] y el comportamiento de las personas dentro de un entorno para observar si en el lugar hay algún riesgo o patrón de comportamiento.

Existe un gran número de aplicaciones de seguimiento que se pueden encontrar en interiores, sin embargo, actualmente se han realizado sistemas de este tipo en escenarios más complejos como por ejemplo entornos naturales o urbanos [2]. La mayoría de los sistemas de seguimiento de personas u objetos móviles se basan en técnicas las cuales detectan el movimiento en imágenes estáticas [2], por lo tanto, al implementar un sistema de seguimiento con cámaras PTZ¹ se está ampliando el campo de visión que se puede obtener, ya que estas por tener movimiento vertical y horizontal, logran abarcar un rango dinámico de mayor visión, realizando así un sistema más eficiente y completo que el que se puede obtener con una cámara estática.

Cada año se adquieren un gran número de sistemas de seguridad los cuales utilizan cámaras con el objetivo de asignar a una persona la labor de estar monitoreando constantemente las escenas de la cámara en tiempo real. El propósito de estas instalaciones es prevenir o detener un incidente indeseable, básicamente con una persona observando un monitor [3]. Esto es un enfoque muy poco realista a la seguridad. Hace 30 años se realizaron experimentos en los Laboratorios Nacionales de Sandia (Sandia National Laboratories) para el Departamento de Energía de Estados Unidos con el fin de poner a prueba la eficacia de una persona cuya tarea era sentarse delante de uno o varios monitores de video por varias horas al día y observar determinados eventos. Estos estudios demostraron que de esta forma, aun cuando se asigna a una persona dedicada y bien intencionada, no se obtendrá un sistema de seguridad eficaz. Después de 20 minutos de observación y evaluación de la pantalla, la atención de la mayoría de los individuos se empieza a perder, hasta llegar a ser por debajo de los niveles aceptables [3]. Por esta razón, se propone utilizar un sistema de seguridad automático el cual está monitoreando continuamente.

¹ PTZ (Pan-Tilt-Zoom). Pan: Movimiento azimut de la cámara. Tilt: Movimiento de inclinación de la cámara.

El objetivo principal de este proyecto es diseñar un sistema de seguimiento de objetos móviles usando una cámara PTZ. El proyecto se divide en dos partes, implementar un algoritmo de seguimiento a partir de filtro de partículas [8] [9] y realizar un piloto con una cámara PTZ para evaluar el seguimiento del objeto móvil. Estas dos etapas se realizan independientemente y a partir del algoritmo de seguimiento se define el tipo de objeto móvil a seguir. Finalmente se implementa un solo algoritmo donde estén las dos etapas iniciales en un solo conjunto para realizar el seguimiento de objetos móviles por medio de una cámara PTZ.

En el presente libro se encuentra el proceso que se llevó a cabo para la realización de este proyecto. Se encuentra la teoría necesaria para comprender este proceso, las especificaciones del proyecto, las pruebas realizadas para la elección del objeto de interés a seguir y de diferentes parámetros del sistema, el desarrollo del proyecto y los resultados obtenidos durante el proceso.

2. MARCO TEÓRICO

2.1 FORMACIÓN Y REPRESENTACIÓN DE UNA IMAGEN

La formación de una imagen tiene tres componentes básicos, los cuales son la iluminación, los modelos de reflexión de las superficies que se visualizan y el proceso de formación de la imagen en la retina de los ojos humanos o en el sensor de la cámara. Inicialmente, las imágenes que se forman son señales análogas bidimensionales. Luego la imagen análoga debe ser muestreada, digitalizada y debe ser representada en un formato apropiado para ser procesada y manipulada por un computador [6].

2.2 RECONOCIMIENTO DE PATRONES EN UNA IMAGEN

La decisión de clasificación de patrones es una técnica donde la clasificación de cierto patrón desconocido se decide basándose en principios determinísticos o estadísticos, donde su diagrama de bloques se muestra en la Figura 1.

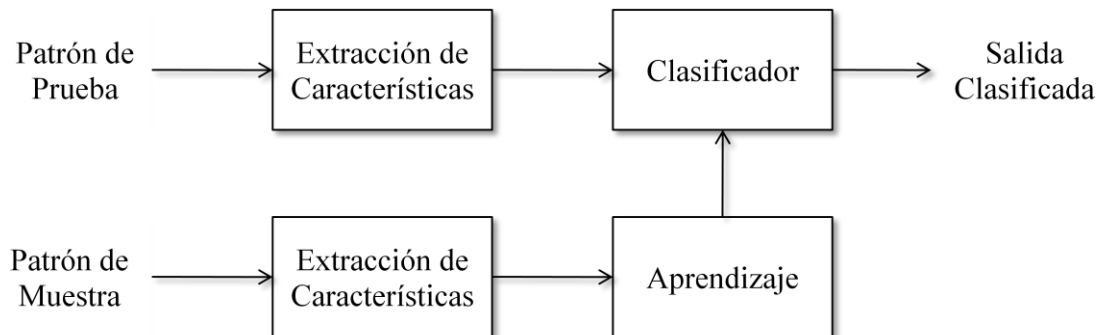


Figura 1 Diagrama de bloques de la decisión teórica de la clasificación de patrones [6]

Estas técnicas son de dos tipos: métodos de clasificación basados en aprendizaje supervisado y métodos de clasificación usando técnicas sin supervisión. Los algoritmos de clasificación supervisada se clasifican en clasificadores paramétricos y no paramétricos. En la clasificación supervisada paramétrica, el clasificador es formado con un conjunto de muestras de patrones conocidos formados etiquetados para estimar los parámetros de cada clase de patrones como media y varianza. Las características de entrada obtenidas de la formación de fase o los clasificadores supervisados se asumen Gaussianos. [6]

2.2.1 CLASIFICACIÓN DE DISTANCIA MÍNIMA

Para medir la similitud y disimilitud entre dos clases de patrones se utilizan las funciones de distancia. Si la distancia entre dos clases de patrones es mínima, entonces la similitud entre estas es grande. El algoritmo encuentra las distancias de una entrada de prueba con las clases representativas. El patrón desconocido es asignado a la clase donde su distancia sea la más pequeña entre las distancias con las otras clases.

Si A y B son dos objetos o patrones, la distancia entre estos se puede calcular o medir con diferentes funciones de distancia. Las diferentes funciones de distancia son la distancia *Minkowski*, *Manhattan* o *City-Block*, *Euclidiana*, *Euclidiana Cuadrada*, *Chebyshev*, *Mahalanobish* y distancia *encerrada*. [6]

La distancia *Minkowski* se define como:

$$D(A, B) = \sqrt[p]{\sum_{i=1}^n |a_i - b_i|^p}$$

La distancia *Manhattan* o *City-Block* se deriva de *Minkowski* cuando $p = 1$, de la siguiente manera:

$$D(A, B) = \sum_{i=1}^n |a_i - b_i|$$

La distancia *Euclidiana* también se deriva de *Minkowski* cuando $p = 2$, de la siguiente manera:

$$D(A, B) = \sqrt{\sum_{i=1}^n |a_i - b_i|^2}$$

La distancia *Euclidiana Cuadrada* es la distancia *Euclidiana* elevada al cuadrado:

$$D(A, B) = \sum_{i=1}^n (a_i - b_i)^2$$

La distancia *Chebyshev* se define de la siguiente manera:

$$D(A, B) = \max_{i=1, \dots, n} |a_i - b_i|$$

2.2.2 ALGORITMOS DE COMPARACIÓN “MATCHING TEMPLATE”

Para realizar el seguimiento de un objeto en particular, se desea conocer la distribución de probabilidad que representa la hipótesis de la localización del objeto de interés [14].

“*Matching Template*” es un proceso de comparación entre imágenes el cual realiza la función de comparar una sección de la imagen actual con una imagen de entrada, recorriendo la sección de la imagen actual en la imagen de entrada, utilizando diferentes métodos [14].

El propósito de este proceso es comparar dos imágenes parecidas para encontrar la similitud o disimilitud entre las imágenes. En este proyecto se compara una región seleccionada en un cuadro x determinado, con una distribución de regiones aleatorias en el siguiente cuadro para encontrar la región más parecida a la región elegida en el cuadro x .

Entre más complejo sea el método utilizado (por ejemplo el del coeficiente de correlación), la precisión entre las comparaciones será mayor [14]. Sin embargo, esta afirmación no se cumple en todos los casos, ya que cualquiera de estos métodos puede ser el más eficiente dependiendo de su aplicación y de lo que se está buscando lograr con el algoritmo.

2.2.2.1 MÉTODO DE DIFERENCIA CUADRADA

$$R_{sq_diff}(x, y) = \sum_{x', y'} [T(x', y') - I(x + x', y + y')]^2$$

Donde T es la imagen sobre la cual se va a comparar, I es la imagen de entrada y R es el resultado [14]. Si la imagen I es de $W \times H$ y la imagen T es de $w \times h$, entonces el resultado es una imagen de $(W - w + 1) \times (H - h + 1)$ [5].

En este método si el resultado es 0 indica que las dos imágenes “encajan” perfectamente, mientras que si el resultado es muy alto las imágenes coinciden poco. Es decir, entre más pequeño sea el resultado la similitud entre las dos imágenes va a ser mayor.

2.2.2.2 MÉTODO DE CORRELACIÓN

$$R_{ccorr}(x, y) = \sum_{x', y'} [T(x', y') \cdot I(x + x', y + y')]^2$$

En este caso, como las imágenes se están multiplicando, entonces si el resultado es menor indica que las imágenes no coinciden o “encajan” y son muy distintas, mientras que si el resultado es alto, las imágenes serán más parecidas.

La correlación entre dos variables es la relación de una variable con respecto a la otra, es decir, si se tienen dos variables, a y b , cuando a aumenta y b también aumenta, entonces a y b están correlacionadas.

2.2.2.3 MÉTODO DE COEFICIENTE DE CORRELACIÓN

$$R_{ccoeff}(x, y) = \sum_{x', y'} [T'(x', y') \cdot I'(x + x', y + y')]^2$$

Donde,

$$T'(x', y') = T(x', y') - \frac{1}{(w \cdot h) \sum_{x'', y''} T(x'', y'')}$$

$$I'(x + x', y + y') = I(x + x', y + y') - \frac{1}{(w \cdot h) \sum_{x'', y''} I(x + x'', y + y'')}$$

En este método si el resultado es 1 entonces las dos imágenes encajan perfectamente, si el resultado es -1 las imágenes desencajan perfectamente y si el resultado es 0 entonces no existe correlación entre las imágenes.

El coeficiente de correlación indica qué tan correlacionadas se encuentran las variables. Si estas están correlacionadas, cuando la variable b aumenta al aumentar a , pero en menor proporción, entonces estas variables tienen un coeficiente de correlación menor que en el caso de que ambas variables aumentaran en la misma proporción. Si a medida que aumenta a , disminuye b , entonces se dice que el coeficiente de correlación entre estas variables es negativo.

El rango de los valores que puede tomar el coeficiente de correlación es de -1 a 1, donde toma el valor de -1 cuando una variable disminuye a medida que la otra aumenta con la misma proporción, y 1 cuando ambas variables aumentan o disminuyen con la misma proporción.

2.3 SEGUIMIENTO DE OBJETOS EN MOVIMIENTO

En general, el seguimiento de objetos se puede realizar de dos maneras, seguimiento basado en reconocimiento y seguimiento basado en movimiento. Un sistema de seguimiento de objetos rápidos se desarrolla basado en técnicas de movimiento predictivo como por ejemplo el filtro de Kalman o filtro de partículas. En el procesamiento de imágenes automatizado basado en sistemas de seguimiento, los objetos de interés se encuentran automáticamente sin intervención humana [6].

Un video o escena dinámica es una secuencia de imágenes o cuadros interrelacionados, donde cada cuadro representa la escena en un determinado instante de tiempo t . Por lo tanto una video puede representarse como función $F(x, y, t)$, donde x y y son las coordenadas espaciales del cuadro actual y t es el tiempo. Para realizar el procesamiento de una escena dinámica se pueden realizar técnicas estáticas a cada cuadro independiente o analizar una secuencia de manera continua [9].

En el caso de seguimiento basado en reconocimiento, se determina un patrón característico del objeto que se desee seguir y este es reconocido en cada imagen de cada uno de los cuadros sucesivos del video y se lleva a cabo el seguimiento usando la información de la posición [6].

Un sistema de seguimiento se puede describir en tres pasos secuenciales: estado inicial, estado de seguimiento y estado de recuperación o restablecimiento. El estado inicial se presenta cuando la cámara está en modo de búsqueda del objeto que se desea seguir. En este estado, el procesamiento se lleva a cabo en la imagen completa del cuadro actual del video. Luego, cuando el objeto de interés a seguir es detectado y localizado el sistema pasa al estado de seguimiento. En este estado se identifica la siguiente posición del objeto a seguir y se crea una historia de esta información. Si el objeto de interés no se encuentra en el área predicha de la ventana, entonces el sistema pasa al siguiente paso que es el paso de recuperación o restablecimiento. En este estado cuando el objeto de interés se pierde el sistema trata de recuperarlo. Si se logra recuperar entonces el sistema vuelve al estado de seguimiento, de lo contrario se queda en el mismo estado

durante un intervalo de tiempo definido. Si el llega a dicho tiempo y no ha encontrado el objeto, entonces vuelve al estado inicial [6].

Para realizar el seguimiento de la trayectoria de un objeto en movimiento se debe tener en cuenta que la posición, la velocidad y la dirección de movimiento de cierto punto dado de un cuadro al siguiente es relativamente invariante o varía muy poco [9].

2.3.1 SEGUIMIENTO DE OBJETOS BASADO EN FILTRO DE PARTÍCULAS

Uno de los objetivos del seguimiento es realizar la predicción y la actualización del estado futuro a partir de una secuencia de observaciones. Un método para realizar esto es el Filtro de Partículas, donde se realiza el proceso de actualización y predicción por medio de la evolución temporal de las partículas como se muestra en la Figura 2. [9]

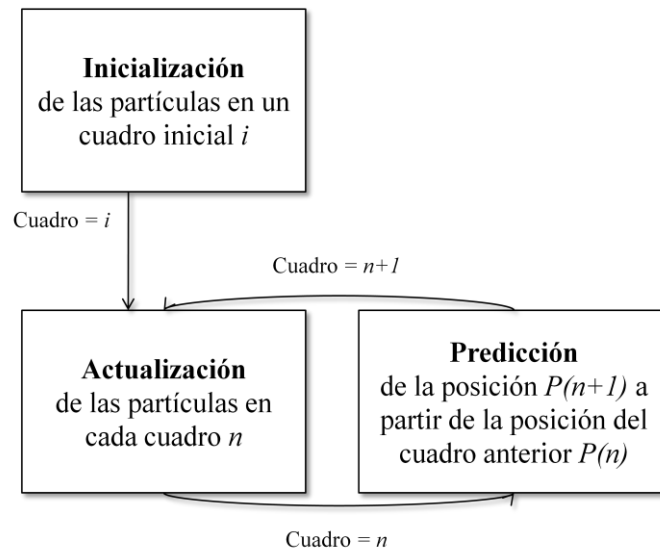


Figura 2 Actualización y Predicción

El filtrado de partículas es un método de procesamiento de imágenes, el cual corresponde a los algoritmos tipo Monte-Carlo [15] donde éste se implementa secuencialmente basado en modelos aleatorios. Este método aproxima las funciones de densidad de probabilidad y su evolución en el tiempo directamente como un número finito de muestras discretas. Estas muestras discretas se denominan *partículas* y están conformadas por un peso y un estado [9]. El enfoque del filtro de partículas es aproximar distribuciones relevantes con diferentes mediciones aleatorias compuestas por partículas aleatorias. Cada partícula tiene diferentes características como posición, tamaño y

forma, ya que esta última debe elegirse para definir las regiones alrededor del origen de cada partícula y puede ser regular o arbitraria [6].

El algoritmo de filtro de partículas consta de los siguientes pasos: Inicialización, re-muestreo y estimación. En la inicialización es donde se crea una población inicial de M partículas usando una función de densidad de probabilidad conocida o aleatoria. En el re-muestreo se elige el rango de la ubicación de las partículas y su tamaño. Finalmente, la estimación dinámica se puede representar como el movimiento del objeto de un cuadro a otro, es decir, en este paso es donde se predice la posición del objeto en el siguiente cuadro con respecto a la posición del objeto en el cuadro actual.

2.4 CÁMARA SONY EVI-D70

La cámara con la que se desarrolla el proyecto es una cámara pan-tilt-zoom de referencia SONY EVI-D70. Esta se puede observar en la Figura 3.



Figura 3 CÁMARA SONY EVI-D70 [4]

El control de esta cámara se realiza mediante un control remoto o desde un computador mediante los comandos VISCA [4]. Para este proyecto interesa el manejo de la cámara mediante los comandos VISCA, ya que el propósito del presente trabajo de grado es realizar un piloto con la

cámara PTZ el cual evalúe el seguimiento del objeto de interés automáticamente sin que el usuario realice manualmente el movimiento de la cámara.

Unas de las principales características que son útiles para el presente proyecto se encuentran en la Tabla 1.

CARACTERÍSTICAS DE LA CÁMARA EVI-D70		
Imágenes de alta resolución	380,000 píxeles	
Alta velocidad de obturación	de 1/1 a 1/10,000 segundos	
Tiene un amplio ángulo de visión (PAN y TILT)	PAN	desde -170 hasta +170 grados
	TILT	desde -30 hasta +90 grados
Posibilidad de invertir la imagen para aplicaciones que requieren que la cámara se ubique en el techo	IMAGE FLIP ON - OFF	
RS-232 y RS-422 externos		
Zoom Óptico de lente	de 1x a 18x	
La calidad de la imagen del zoom digital es mejorada	DSP (<i>Digital Signal Processor</i>)	
Infrarrojo opcional	IF ON – OFF	

Tabla 1. Características principales de la cámara EVI_D70 [4]

2.4.1 ZOOM

El zoom de esta cámara puede ser zoom óptico y digital. El zoom óptico tiene un alcance de 18x y el digital tiene un alcance de 216x, sin embargo cuando se llega a un alcance de 12x el zoom digital empieza a ser utilizado, y la resolución se pierde. El zoom digital funciona ampliando el centro del sujeto expandiéndolo vertical y horizontalmente [4].

2.4.2 LISTA DE COMANDOS

El protocolo que se utiliza para enviar y recibir los datos es el protocolo VISCA. El controlador de estos comandos es el computador y el que recibe los comandos es el dispositivo periférico EVI-D70. La comunicación entre estos dos se realiza mediante RS-232 y RS-422.

Los parámetros de RS-232 y RS-422 se muestran en la Tabla 2.

PARÁMETROS DE RS-232 Y RS-422	
Velocidad de comunicación	9,600 bps / 38,400 bps
Data bits	8
Start bit	1
Stop bit	1
Paridad	Ninguno
Control de flujo	Ninguno

Tabla 2 parámetros de RS-232 y RS-422 [4]

La cámara EVI-D70 tiene 2 buffers para comandos, por lo tanto pueden ser recibidos hasta 2 comandos incluyendo los que están siendo ejecutados. Cuando recibe un comando, notifica cual buffer de comandos fue utilizado [4].

2.4.3 PAN - TILT

Las características de los movimientos de PAN y TILT se muestran en la

PAN – TILT			
Velocidad máxima	PAN	100 grados por segundo	
	TILT	90 grados por segundo	
Ángulo de visión	PAN	desde -170 hasta +170 grados	Con incrementos de 0.075 grados
	TILT	desde -30 hasta +90 grados	

Tabla 3 Movimientos de PAN y TILT [4]

3. ESPECIFICACIONES

3.1 DIAGRAMA GENERAL

En la Figura 4 se puede observar el esquema general del funcionamiento del proyecto.

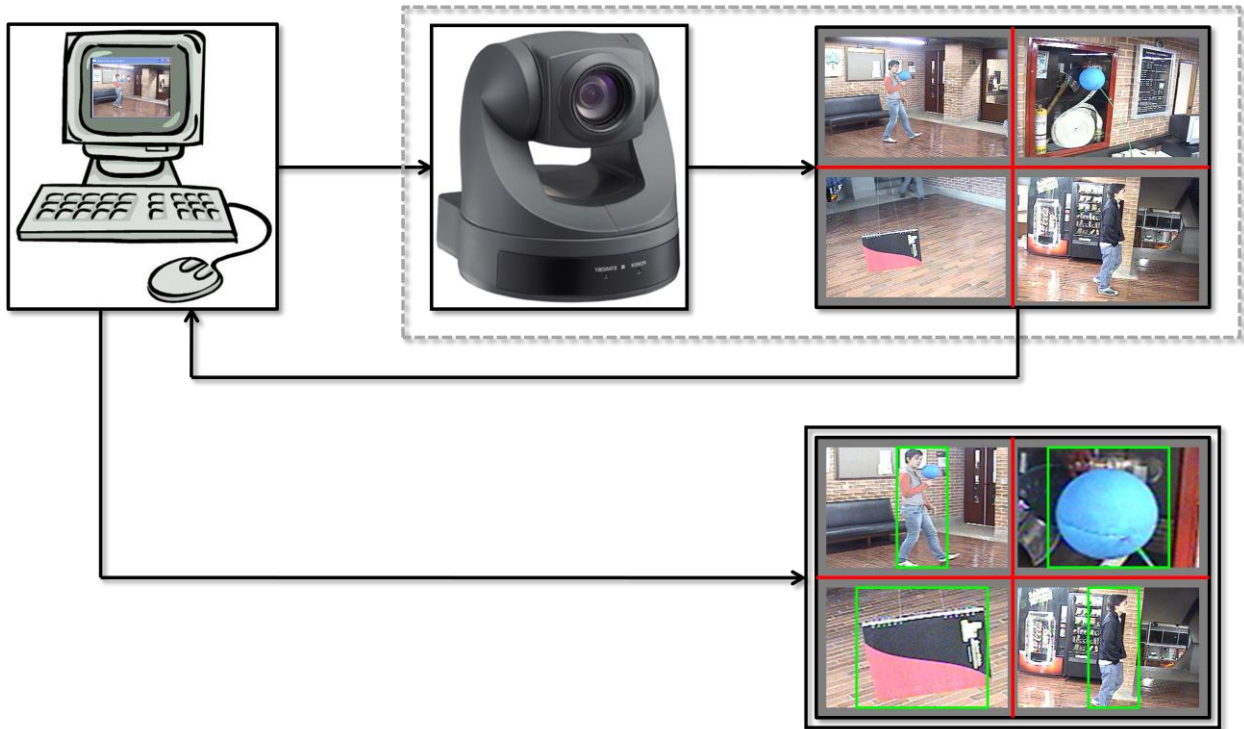


Figura 4 Diagrama de Bloques general [16],[4]

3.2 ESPECIFICACIONES GENERALES

- Entradas y salidas de la cámara

Entradas	Órdenes de movimiento pan, tilt y zoom (provenientes del computador)
Salidas	Video1 (entrada al computador)

- Entradas y salidas del computador

Entradas	Video1 (salida de la cámara)
	Información de las características del video
Salidas	Órdenes de movimiento pan, tilt y zoom (se envían a la cámara)
	Video donde se observa el seguimiento de objeto de interés

El seguimiento se realiza basándose en el método de Filtro de Partículas en tiempo real, donde el objeto de interés a seguir es seleccionado manualmente por el usuario en cualquier momento donde desee empezar el seguimiento. El objeto de interés se sigue de tal forma que este siempre se encuentre ubicado en el centro de la pantalla. A medida que se va haciendo el seguimiento se realiza un acercamiento al objeto de interés, de tal manera que este ocupe una gran porción del tamaño de la ventana, para poder ser mejor detallado dependiendo de su aplicación.

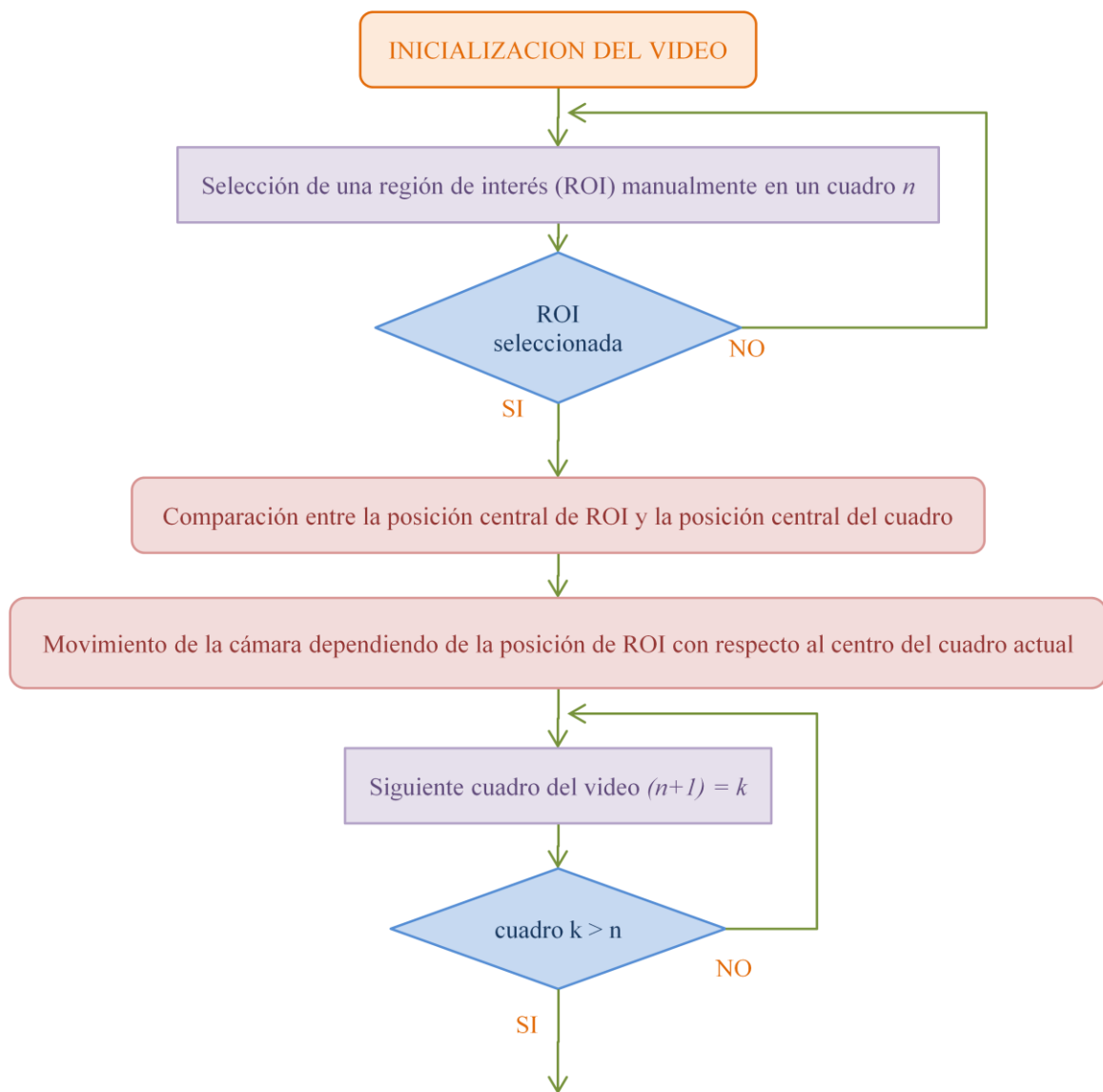
El objeto de interés se determina de acuerdo a las pruebas realizadas, en las cuales se encuentren los resultados más favorables. Este objeto puede ser rígido, donde este permanece relativamente constante a través del tiempo y no presenta variaciones muy grandes durante el seguimiento, o también puede ser flexible, donde este varía significativamente a través del tiempo. Un ejemplo de un objeto flexible puede ser una persona, donde su posición puede cambiar significativamente al caminar.

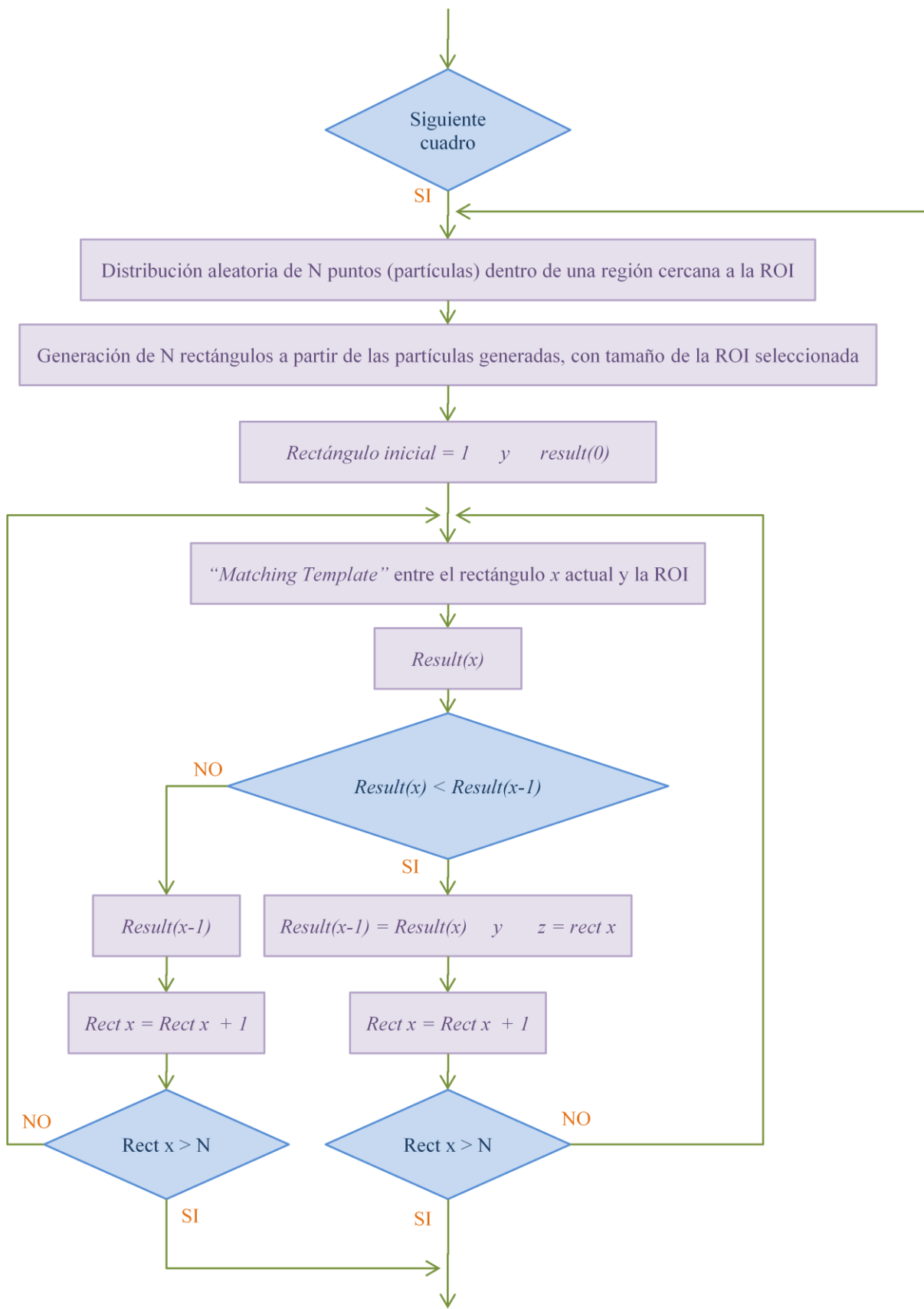
3.3 DESCRIPCIÓN

El proyecto se realiza con la herramienta OpenCV aplicando visión artificial. Este consta de tres fases, la primera consiste en implementar un algoritmo de seguimiento del objeto de interés. Se requiere ubicar inicialmente el objeto en escena ya que a partir de esta ubicación es que se realiza el seguimiento, el cual se basa en el método de filtro de partículas. Este seguimiento lleva un registro espacial temporal de la trayectoria del objeto. La segunda fase consiste en usar este registro espacial para variar los parámetros de la cámara PTZ de tal manera que se mantenga el objeto móvil dentro de la escena de la imagen mientras este se mueve. La tercera fase es definir un objeto de interés a seguir. Este objeto se define a partir de las pruebas realizadas basándose en el comportamiento observado en el algoritmo para cada uno de los posibles objetos a seguir.

Para realizar el seguimiento primero se selecciona manualmente una región de interés a seguir en la escena del video, luego en el siguiente cuadro, aplicando el Filtro de Partículas, se crea una distribución de rectángulos (partículas) para realizar la predicción de la posición del objeto en el siguiente instante de tiempo. Cuando estas regiones están creadas entonces se compara cada una de estas con la región seleccionada inicialmente y se elige el rectángulo que tenga mayor similitud con la región de interés inicial. Luego, el rectángulo elegido se convierte en la región de interés a seguir y en el siguiente cuadro se vuelve a crear otra distribución de cuadros y se realiza sucesivamente este proceso cuadro a cuadro para ir encontrando los rectángulos con mayor similitud y realizar el seguimiento del objeto de interés a seguir.

3.4 DIAGRAMA DE FLUJO DEL MÉTODO DE FILTRO DE PARTÍCULAS





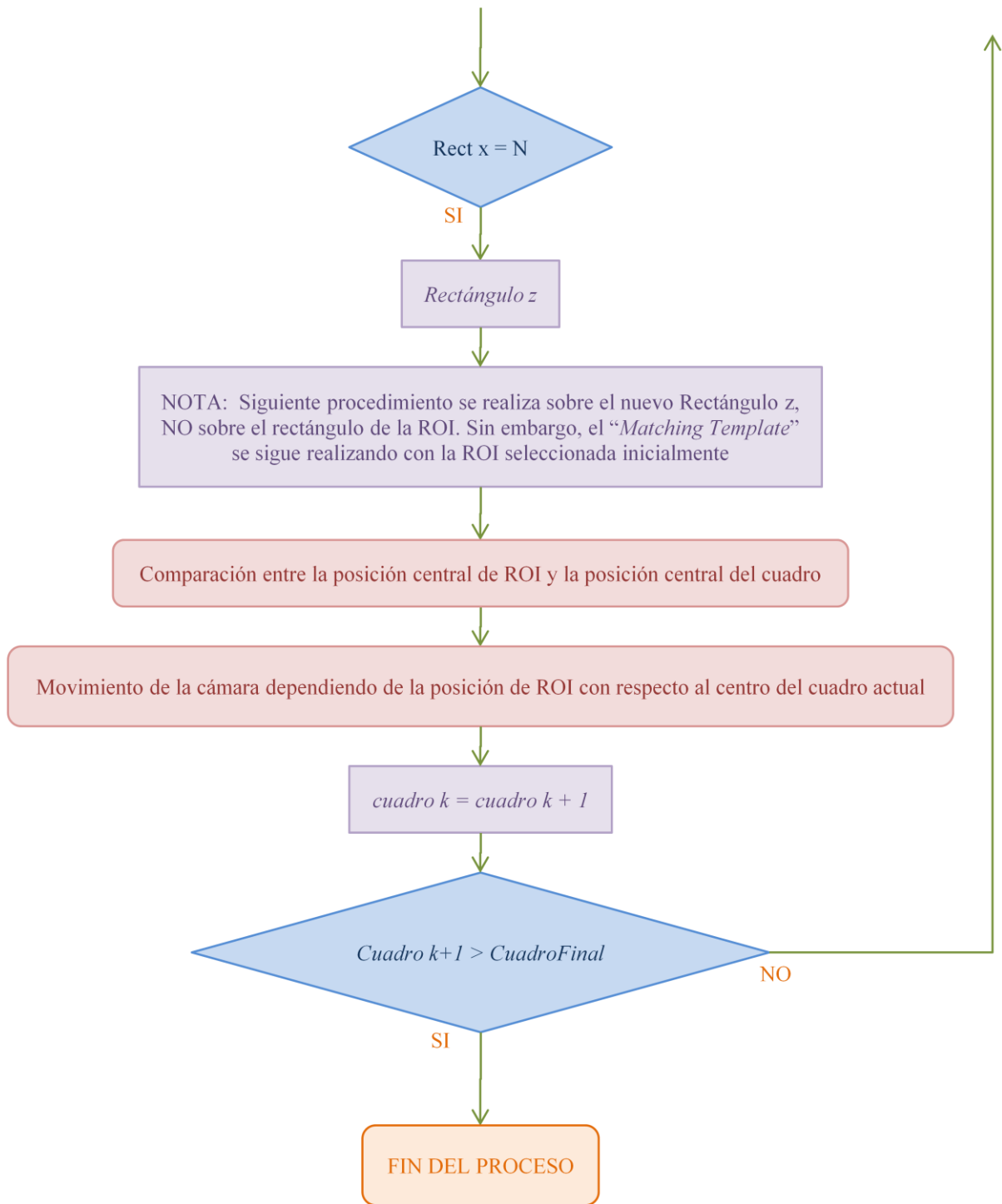


Figura 5 Diagrama de flujo del Filtro de Partículas

3.5 DIAGRAMA DE FLUJO DEL MOVIMIENTO DE LA CÁMARA

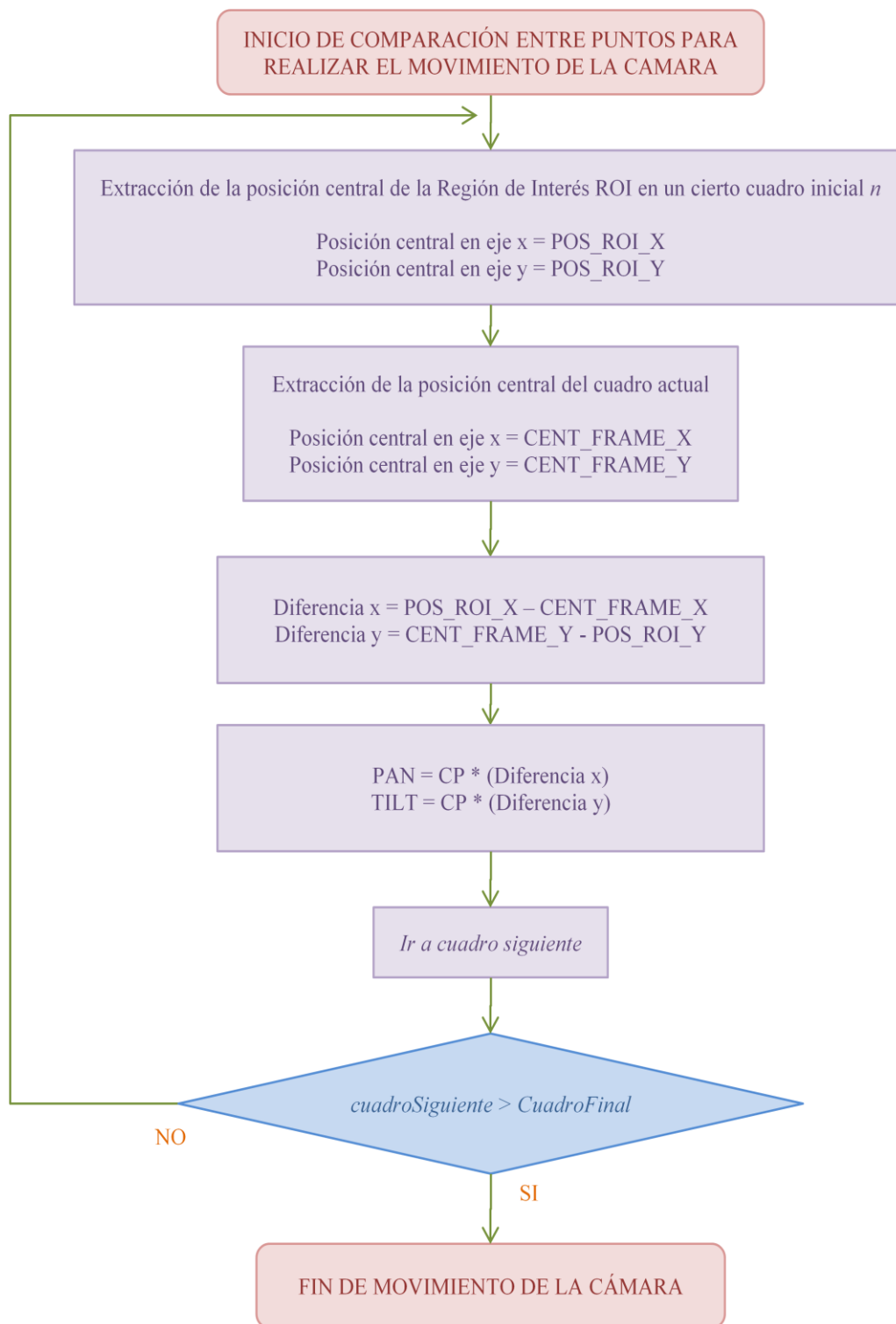


Figura 6 Diagrama de flujo del movimiento de la cámara

4. DESARROLLOS

El proyecto se divide en dos etapas, la implementación de un algoritmo de seguimiento basado en el método de Filtro de Partículas y el desarrollo de un piloto con una cámara PTZ para evaluar el seguimiento de un objeto de interés en movimiento. Estas se realizan independientemente y a partir del algoritmo de seguimiento se define el tipo de objeto móvil a seguir. Finalmente, se implementa un solo algoritmo, en el cual se complementan el algoritmo de seguimiento y el algoritmo del piloto de la cámara PTZ para realizar el seguimiento de tal manera que se obtenga un mayor ángulo de visión que además pueda moverse para que el objeto de interés sea seguido de tal manera que este siempre se encuentre ubicado en el centro de la pantalla.

Ambos algoritmos, tanto el de seguimiento basado en el método de Filtro de Partículas, como el del piloto de la cámara PTZ se implementan bajo C usando bibliotecas nativas de Windows para el manejo del puerto serial y OpenCV para realizar el procesamiento de la imagen.

4.1 DESARROLLO DEL ALGORITMO DE FILTRO DE PARTÍCULAS

El desarrollo general del algoritmo de seguimiento basado en el Filtro de Partículas se encuentra en el diagrama de flujo correspondiente a este proceso ilustrado en la Figura 5, mientras que en esta sección se describe más detalladamente. El código realizado se encuentra en el ANEXO D y ANEXO E.

El primer paso para empezar con el seguimiento es iniciar el video. Después de que este se inicia, se procede a seleccionar manualmente la región de interés en el momento que el usuario desee realizar el seguimiento. Esto se ilustra en la Figura 7.



Figura 7 Selección Manual de la Región de Interés

Al tener la región de interés, esta es guardada en otra imagen diferente para posteriormente realizar las comparaciones correspondientes y el video cambia de un cuadro al siguiente. Esto se muestra en la Figura 8.

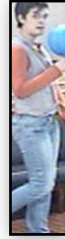


Figura 8 Región de Interés ROI seleccionada

Luego, el video se mueve al siguiente cuadro. En este nuevo cuadro ($k = n + 1$) se toma un punto de la posición de la región de interés (este punto puede elegirse en cualquiera de las esquinas de la región, así como también en el centro de la región) y a partir de este punto se crea una ventana o región donde su tamaño puede cambiarse según las preferencias, y su centro es la posición del punto de la región de interés con respecto a la imagen o el cuadro total. La creación de la ventana se observa en la Figura 9.



Figura 9 Ventana con un largo y un ancho determinado

En esta ventana se crea una distribución de N partículas (puntos) aleatorias, donde la cantidad de partículas N también se puede cambiar según las preferencias. A partir de cada una de estas partículas se generan N rectángulos del mismo tamaño de la región de interés, donde la posición del punto elegido anteriormente es el mismo punto donde se generan los rectángulos, es decir, si el punto elegido de la región de interés es el punto de la esquina superior izquierda entonces a

partir este punto se crea una ventana de un ancho y un largo determinado, donde el punto es el centro de la ventana. Dentro de esta ventana se crea una distribución de N partículas aleatorias y cada una de estas partículas es la esquina superior izquierda de los N rectángulos creados como se ilustra en la Figura 10.



Figura 10 Generación de los rectángulos aleatorios en el cuadro $k = n+1$

Después de generar los N rectángulos se almacena temporalmente en una imagen diferente cada una de las regiones encerradas por cada uno de los rectángulos. Se recuerda que la región de interés seleccionada es una imagen correspondiente al cuadro n y que la distribución de partículas se genera sobre el cuadro $k = n + 1$. Luego se realiza una comparación de imágenes entre la región de interés inicial y cada uno de los N rectángulos generados. Esta comparación se lleva a cabo mediante el algoritmo de “*Matching Template*”.

Como se menciona en la sección 2.2.2 el algoritmo de “*Matching Template*” consta de tres métodos: Diferencia cuadrada, Correlación y Coeficiente de correlación. El algoritmo de seguimiento planteado se trabaja con el método de diferencia cuadrada y el método del coeficiente de correlación. El propósito de esto es comparar el comportamiento de ambos métodos y elegir uno con el cual se obtengan mejores resultados de seguimiento.

Dependiendo del método utilizado se debe realizar un cambio en el algoritmo para cada uno de los métodos, ya que como se menciona en la sección 2.2.2, el resultado de la comparación entre dos imágenes con el método de diferencia cuadrada muestra mayor similitud cuando el resultado de la diferencia cuadrada es mínimo. Sin embargo, cuando se utiliza el método de coeficiente de correlación sucede todo lo contrario. El resultado de la comparación entre dos imágenes en este caso, muestra mayor similitud cuando el resultado del coeficiente de correlación es máximo. El rango posible que puede obtenerse utilizando este método es de -1 a 1, donde el valor 1

representa la máxima similitud (imágenes iguales) y el valor -1 representa la máxima disimilitud. Por lo tanto, cuando se comparan dos imágenes y se habla de similitud, se está haciendo énfasis a ambos algoritmos. Simplemente, cuando el método es diferencia cuadrada, la mayor similitud se obtiene con el resultado mínimo y cuando el método es el coeficiente de correlación, la mayor similitud se obtiene con el resultado máximo.

Retomando el algoritmo de seguimiento, al realizar la comparación entre la imagen de la región de interés inicial y cada uno de los N rectángulos generados por medio del algoritmo de “*Matching Template*”, se encuentra el resultado de cada una de estas comparaciones, donde estos a la vez son comparados entre sí para encontrar el rectángulo donde se obtenga mayor similitud. Luego, después de encontrar el rectángulo con mayor similitud, este es el rectángulo elegido sobre el cual se realiza la distribución de partículas y rectángulos en el siguiente cuadro. Luego, cuando el video avanza un cuadro se genera nuevamente la distribución de partículas y se realiza el mismo procedimiento anterior hasta llegar al cuadro final del video o que el usuario decida terminar el programa.

4.2 DESARROLLO DEL ALGORITMO DEL PILOTO DE LA CÁMARA PTZ

El desarrollo general del algoritmo del piloto de la cámara PTZ se encuentra en el diagrama de flujo correspondiente a este proceso ilustrado en la Figura 6, mientras que en esta sección se describe más detalladamente. El código realizado se encuentra en el ANEXO F.

Para realizar el movimiento de la cámara de tal forma que el objeto de interés se encuentre ubicado siempre en el centro de la pantalla, lo primero que se hace es extraer la posición central de la región de interés, si se encuentra en el primer cuadro, o sino la posición central del rectángulo elegido por tener la mayor similitud, y la posición central del cuadro actual con respecto al eje x y y . Ahora, conociendo que al compilar un video se obtienen imágenes cuyas coordenadas son como se ilustran en la Figura 11, se realiza la diferencia en x y en y entre el punto medio de la región elegida y el punto medio del cuadro actual de la siguiente manera:

$$\textit{Diferencia en } x = \textit{Punto central Región en } X - \textit{Punto central Cuadro en } X$$

$$\textit{Diferencia en } y = \textit{Punto central Cuadro en } Y - \textit{Punto central Región en } Y$$

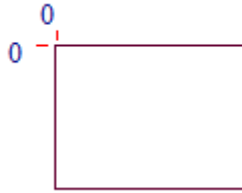


Figura 11 Coordenadas de una imagen

La posición inicial de la cámara se encuentra en cierto valor, el cual puede ser cero. Luego el algoritmo de seguimiento realiza su función y cuando se encuentra una nueva región de interés o un rectángulo elegido, en cada uno de los cuadros entonces a la posición actual de la cámara se le suma un valor que depende de la *diferencia en x* y *diferencia en y*. Este valor se define como una constante, la cual puede ser cambiada por el usuario dependiendo de lo que se desee lograr multiplicada por la *diferencia en x* y *diferencia en y*. Es decir, a la posición actual del ángulo PAN se le incrementa un valor de una constante multiplicada por la *diferencia en x*, y a la posición actual del ángulo TILT se le incremente el valor de una constante, la cual puede ser igual, menor o mayor, a la constante definida para PAN, y esta constante es multiplicada por la *diferencia en y*. Esto se puede mostrar cómo:

$$PAN = PAN_{actual} + diferencia\ en\ x$$

$$TILT = TILT_{actual} + diferencia\ en\ y$$

La *diferencia en x* y *diferencia en y* se definen de esta manera de acuerdo a las cuatro opciones que se presentan, como se observa en la Figura 12.

CONDICIÓN	CARACTERÍSTICA O DESCRIPCIÓN
<i>diferencia en x = positiva</i>	PAN: La cámara se mueve hacia la derecha
<i>diferencia en x = negativa</i>	PAN: La cámara se mueve hacia la izquierda
<i>diferencia en y = positiva</i>	TILT: La cámara se mueve hacia arriba
<i>diferencia en y = negativa</i>	TILT: La cámara se mueve hacia abajo

Figura 12 Distancia en x y en y entre el punto central de la región y el punto central del cuadro

Cuando la diferencia es negativa entonces los ángulos de PAN o TILT se mueven en direcciones opuestas ya que no se estaría sumando o incrementando un valor, sino que este se le está restando a la posición actual.

Este algoritmo permite que el objeto de interés a seguir se mantenga en todo momento en el centro de la pantalla debido a que no se permite que el punto central del rectángulo se salga del centro de la imagen, ya que cuando esta posición central del rectángulo se encuentra en un valor diferente a la posición central del cuadro, inmediatamente se mueve la cámara con respecto al ángulo PAN y/o al ángulo TILT, dependiendo de la posición, de tal manera que ese punto se desplace al punto central del cuadro.

4.3 PROTOCOLO DE PRUEBAS

Para la realización de las pruebas se toman videos de 3 tipos de objetos de interés a seguir diferentes. Para cada uno de éstos se graba un video con fondo controlado y un video con fondo no controlado. La definición de los tipos de objeto de interés a seguir se muestra en la Tabla 4.

POSIBLES OBJETOS DE INTERÉS A SEGUIR	
Objeto rígido sin aristas	Objeto en forma esférica de un solo color donde su posición no cambia a través de los cuadros
Objeto rígido con aristas	Objeto en forma rectangular con más de un color donde su posición puede varias a través de los cuadros
Objeto flexible	Este objeto puede ser una persona con múltiples colores donde su posición puede variar significativamente a través de los cuadros

Tabla 4 Posibles objetos de interés a seguir

La descripción de los tipos de fondo controlado y no controlado se observan en la Tabla 5

FONDOS	
Fondo Controlado	El fondo controlado es un fondo totalmente verde, donde el objeto de interés a seguir puede diferenciarse fácilmente con el fondo
Fondo NO Controlado	El fondo no controlado es un fondo cualquiera donde se pueden encontrar problemas con los objetos que hacen parte del fondo ya que en el éste se pueden encontrar diferentes “obstáculos” los cuales pueden confundir el objeto que se desea seguir con el fondo

Tabla 5 Fondos

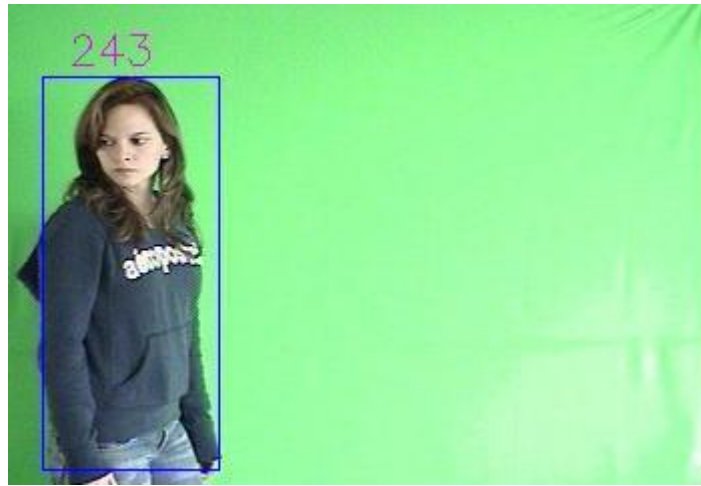
Para cada uno de los videos se realizan diferentes pruebas donde la ventana de la distribución de partículas y la cantidad de partículas utilizadas varían. Esto se realiza con el propósito de compararlos, analizarlos y elegir un tamaño de ventana y cantidad de partículas donde se obtenga un resultado bueno, sin que el algoritmo de seguimiento se vuelva demasiado lento.

En cada video se realiza el seguimiento manualmente mediante el algoritmo del ANEXO G donde manualmente se va seleccionando algún punto en particular del objeto de interés y evaluar el seguimiento con mayor precisión. El propósito de realizar esto es comparar posteriormente estos videos con los videos del seguimiento automático para cada uno de los métodos. El algoritmo del seguimiento automático para cada uno de los métodos se encuentra en el ANEXO H y en el ANEXO I. Para cada una de estas comparaciones se realiza la gráfica de seguimiento de los puntos encontrados manualmente y automáticamente. A partir de estas dos gráficas se encuentra la distancia euclidiana entre cada uno de los puntos en cada cuadro. Esta distancia también se grafica por aparte para observar el error entre los puntos de cada caso. Finalmente se encuentra el promedio de las distancias para cada caso y se observa el comportamiento para cada uno y se observa en cual se encuentran mejores resultados.

5. ANÁLISIS DE RESULTADOS

Para las pruebas mencionadas en el ítem anterior se obtienen los siguientes resultados:

VIDEO PERSONA FONDO CONTROLADO



Método coeficiente de correlación

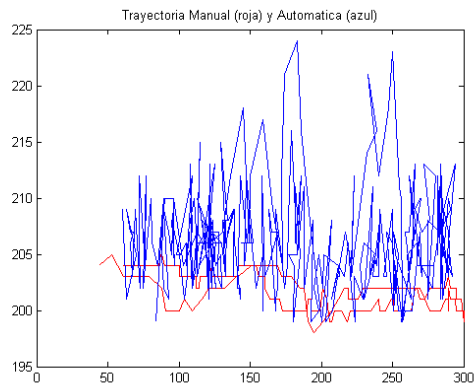


Figura 13 Grafica comparación puntos Manuales y Automáticos

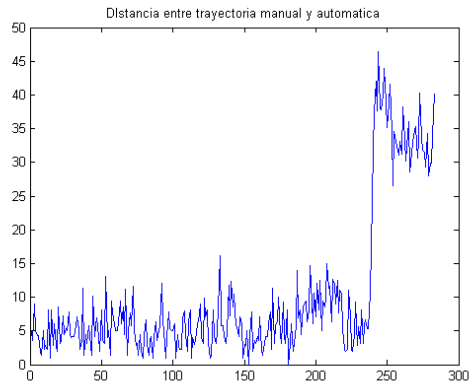
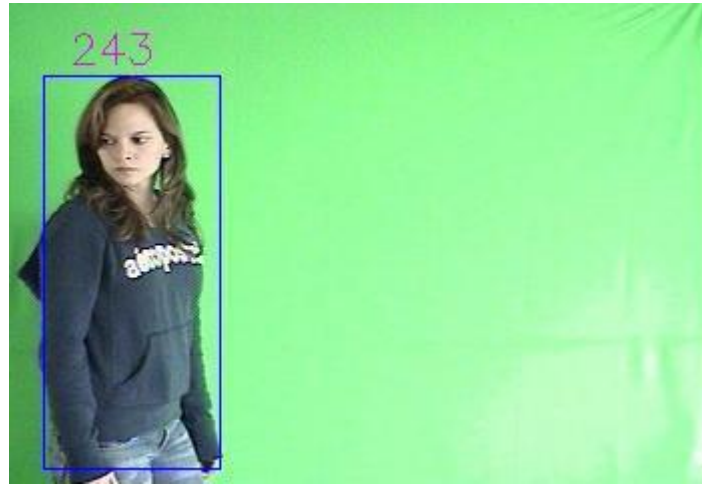


Figura 14 Distancia Euclidiana entre puntos Manuales y Automáticos

Para este video se puede observar que el algoritmo realiza un seguimiento adecuado a la persona debido a que en la gráfica de la distancia euclidiana se observa un error pequeño y constante la mayoría de los cuadros. Al final de la trayectoria se ve que en esa parte la persona se trata de perder, pero no alcanza a perderse.

VIDEO PERSONA FONDO CONTROLADO



Método de diferencia cuadrada

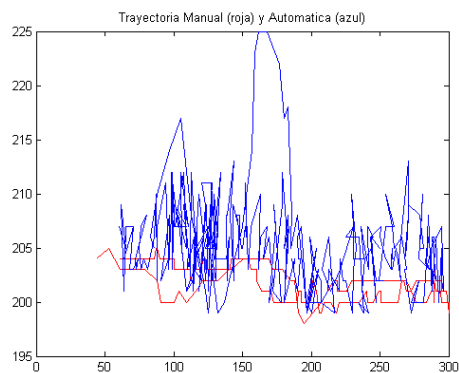


Figura 15 Grafica comparación puntos Manuales y Automáticos

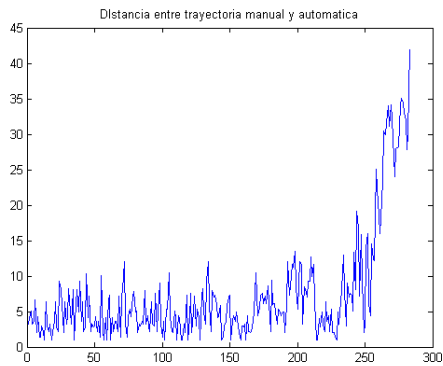


Figura 16 Distancia Euclidiana entre puntos Manuales y Automáticos

En este video también se realiza un seguimiento sin errores. En este el error varía un poco más que con el anterior, pero esto puede deberse al tamaño de la ventana ya que en este caso el tamaño de la ventana es mayor, y cuando la ventana de la distribución de puntos es grande, los rectángulos generados son más dispersos, y pueden existir mayores variaciones en la distancia.

VIDEO OBJETO RIGIDO FONDO NO CONTROLADO



Método de coeficiente de correlación

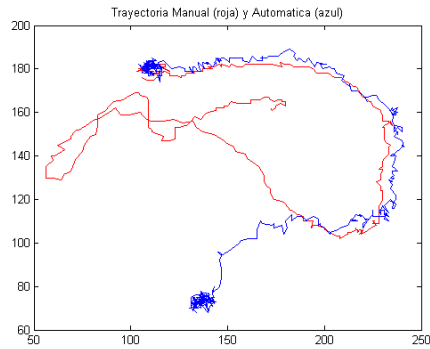


Figura 17 Grafica comparación puntos Manuales y Automáticos

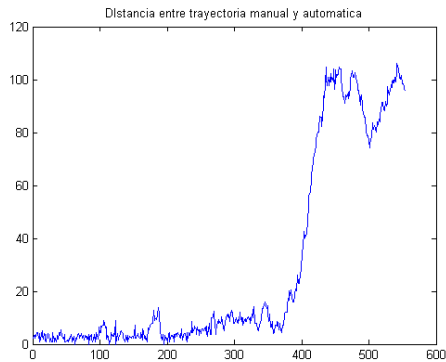


Figura 18 Distancia Euclidiana entre puntos Manuales y Automáticos

En este video realiza un seguimiento con errores. Inicialmente, el objeto se mueve con una velocidad lenta y el algoritmo lo sigue correctamente, sin embargo, cuando este objeto empieza a aumentar la velocidad, como se tiene una ventana pequeña, entonces los rectángulos generados no alcanzan a tomar la totalidad del objeto cuando este se mueve rápidamente.

VIDEO OBJETO RIGIDO FONDO NO CONTROLADO



Método de diferencia cuadrada

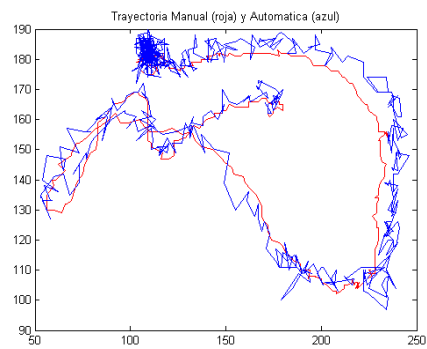


Figura 19 Grafica comparación puntos Manuales y Automáticos

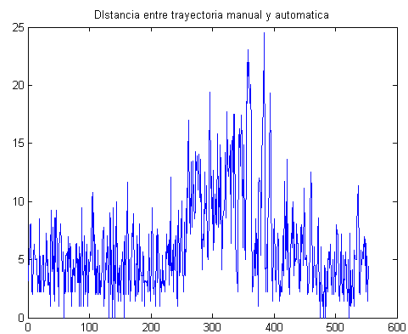


Figura 20 Distancia Euclidiana entre puntos Manuales y Automáticos

En este video se observa un seguimiento correcto. En la gráfica de la trayectoria Manual y Automática se observa que el objeto en ningún momento se pierde y el error entre ambos puntos no es muy grande.

VIDEO PERSONA COMPLETA FONDO NO CONTROLADO



Método de coeficiente de correlación

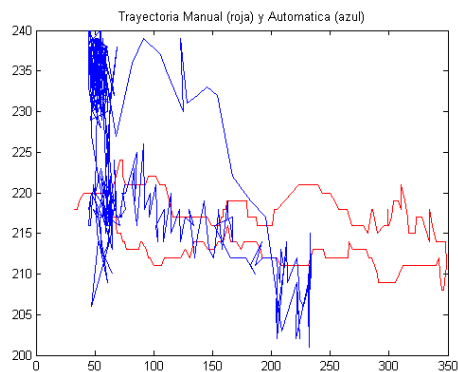


Figura 21 Grafica comparación puntos Manuales y Automáticos

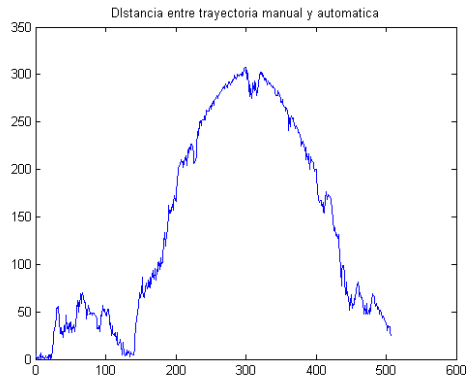


Figura 22 Distancia Euclidiana entre puntos Manuales y Automáticos

Este video muestra un mal seguimiento debido a que la persona es seguida por un intervalo de tiempo muy corto y luego se pierde completamente. Cuando la persona se desplaza de un lado a otro y se devuelve, los rectángulos generados siempre se quedan en una sección de la imagen. Al final de la gráfica de la distancia euclidiana se observa que el error entre los puntos disminuye, pero esta disminución de error se debe a que la persona se devuelve al mismo lugar y como la distribución se queda en ese lugar, entonces la distancia disminuye, pero en realidad no la está siguiendo. Este error se debe a que la ropa de la persona en movimiento tiene un color muy similar a los colores del fondo, por esta razón es que puede perderse, ya que el algoritmo toma partes del fondo como si fueran de la persona.

VIDEO PERSONA COMPLETA FONDO NO CONTROLADO



Método de diferencia cuadrada

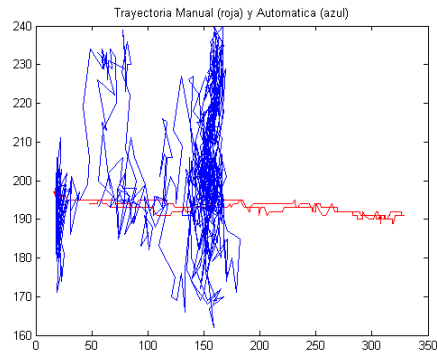


Figura 23 Grafica comparación puntos Manuales y Automáticos

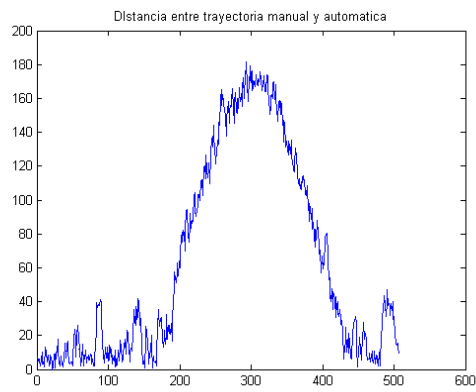


Figura 24 Distancia Euclidiana entre puntos Manuales y Automáticos

En este video se observa el mismo comportamiento anterior, aunque este mejora un poco ya que mantiene el seguimiento por un intervalo de tiempo un poco más largo. El problema de este caso es el video que no es bueno. Si se sigue una persona de tal forma que los colores de la persona contrasten con el fondo el algoritmo funciona correctamente.

VIDEO MEDIA PERSONA FONDO NO CONTROLADO



Método de coeficiente de correlación

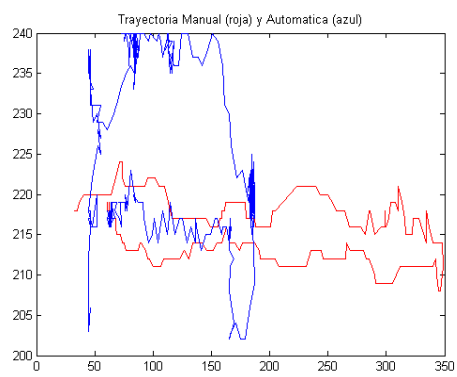


Figura 25 Grafica comparación puntos Manuales y Automáticos

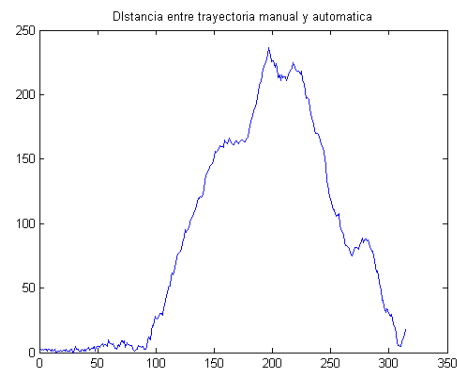


Figura 26 Distancia Euclidiana entre puntos Manuales y Automáticos

En este video es el mismo a se obtienen resultados muy similares a los anteriores, ya que la persona es la misma con la misma ropa la cual no contrasta con el fondo.

VIDEO MEDIA PERSONA FONDO NO CONTROLADO



Método de diferencia cuadrada

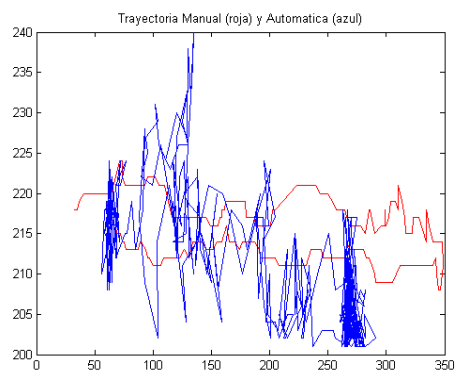


Figura 27 Grafica comparación puntos Manuales y Automáticos

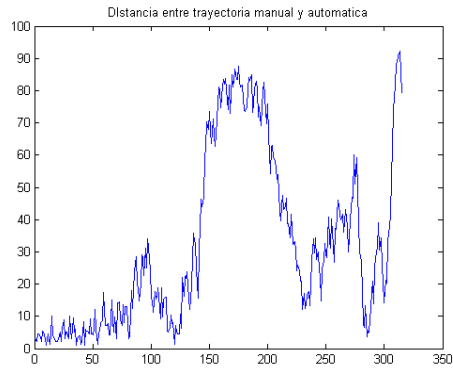


Figura 28 Distancia Euclidiana entre puntos Manuales y Automáticos

Este video tiene resultados mejores que en los anteriores ya que la persona es seguida por más tiempo. Esto se debe a que la ventana es mayor y se toma una mayor cantidad de partículas.

VIDEO OBJETO RIGIDO CON ARISTAS FONDO NO CONTROLADO



Método de coeficiente de correlación

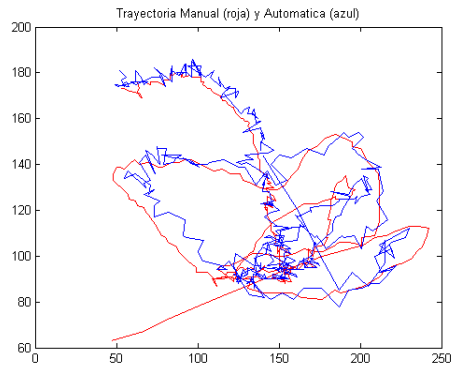


Figura 29 Grafica comparación puntos Manuales y Automáticos

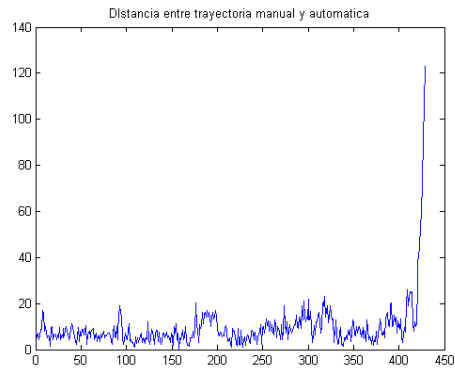


Figura 30 Distancia Euclidiana entre puntos Manuales y Automáticos

En este video se obtiene un seguimiento correcto, ya que el error entre los puntos tomados no es muy grande, y además el objeto no se pierde sino hasta el final del video que se mueve bruscamente para terminar el video.

VIDEO OBJETO RIGIDO CON ARISTAS FONDO NO CONTROLADO



Método de distancia cuadrada

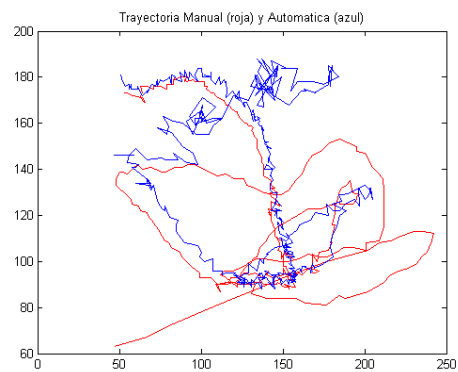


Figura 31 Grafica comparación puntos Manuales y Automáticos

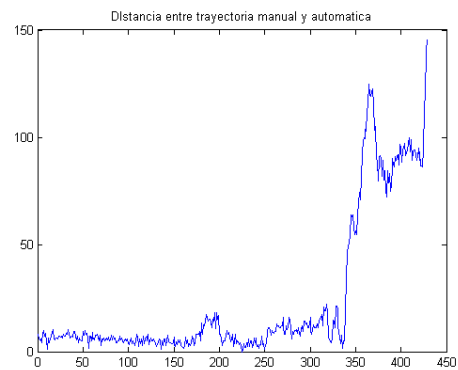
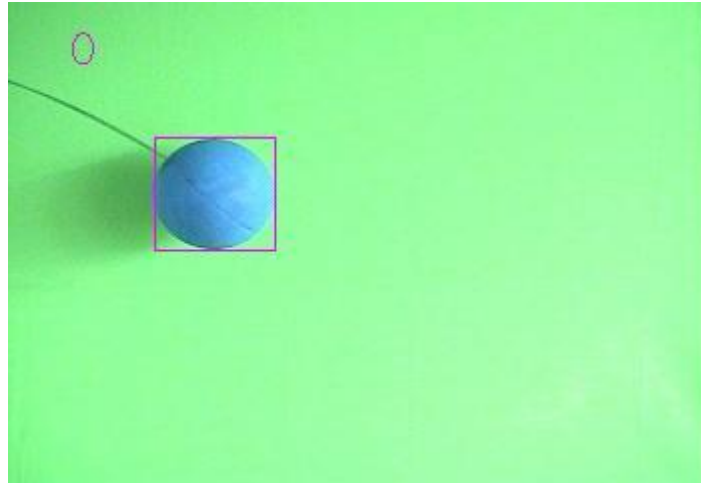


Figura 32 Distancia Euclidiana entre puntos Manuales y Automáticos

Este video es el mismo anterior, y se obtienen peores resultados ya que el error deja de ser casi constante en un punto anterior del caso anterior

VIDEO OBJETO RIGIDO SIN ARISTAS FONDO CONTROLADO



Método de coeficiente de correlación

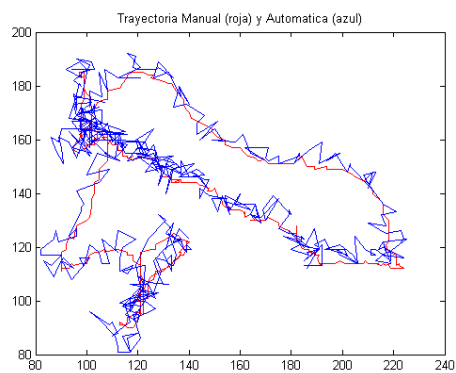


Figura 33 Grafica comparación puntos Manuales y Automáticos

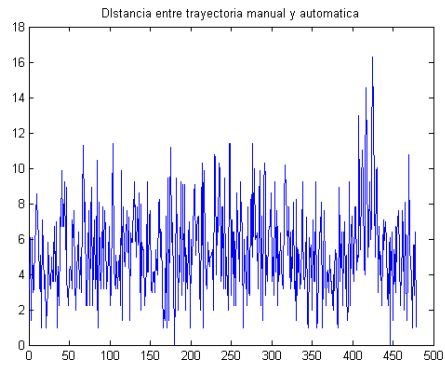
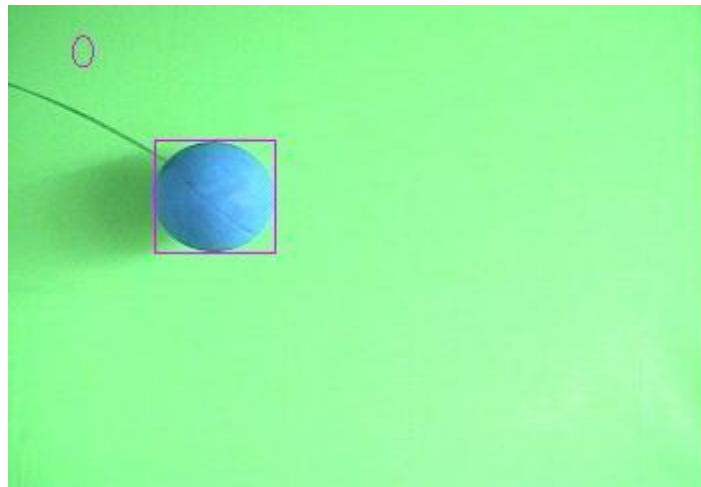


Figura 34 Distancia Euclidiana entre puntos Manuales y Automáticos

Este es un seguimiento muy bueno debido a que el error en ningún momento crece bruscamente, sino que todo el tiempo permanece en un rango de valores cercanos.

VIDEO OBJETO RIGIDO SIN ARISTAS FONDO CONTROLADO



Método de diferencia cuadrada

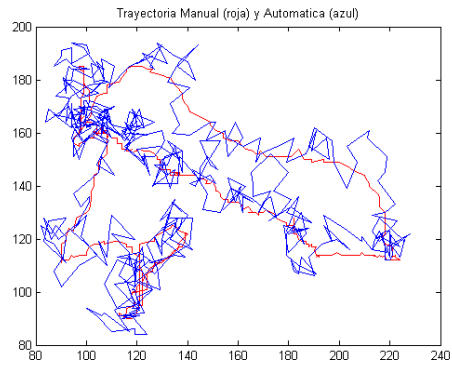


Figura 35 Grafica comparación puntos Manuales y Automáticos

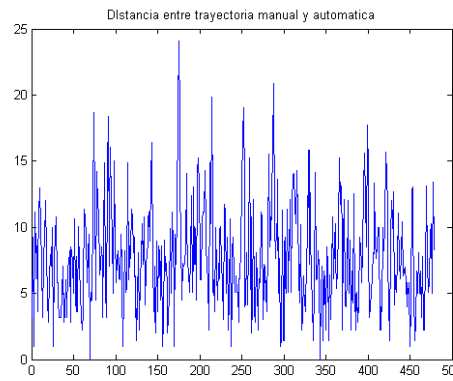


Figura 36 Distancia Euclidiana entre puntos Manuales y Automáticos

Este también es un seguimiento muy bueno, el objeto en ningún momento se pierde. Este video es el mismo anterior, sin embargo la distancia aumenta porque se toma una ventana más grande y los rectángulos se dispersan más, por esto se ven puntos un poco más lejanos, sin representar un mal seguimiento sino una distribución mas distante.

VIDEO OBJETO RIGIDO CON ARISTAS FONDO CONTROLADO



Método d coeficiente de correlación

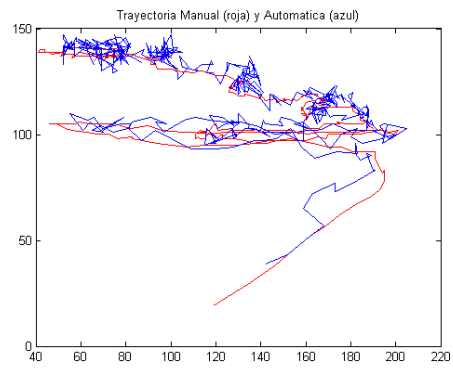


Figura 37 Grafica comparación puntos Manuales y Automáticos

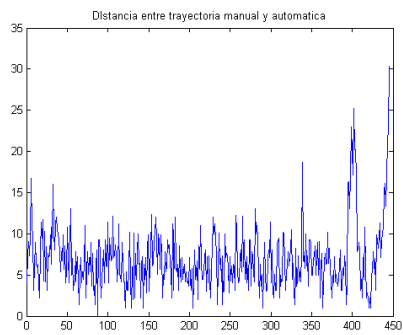


Figura 38 Distancia Euclidiana entre puntos Manuales y Automáticos

Igual que algunos anteriores, en este se obtiene un seguimiento adecuado por las mismas razones de error relativamente constante y pequeño.

VIDEO OBJETO RIGIDO CON ARISTAS FONDO CONTROLADO



Método de diferencia mínima

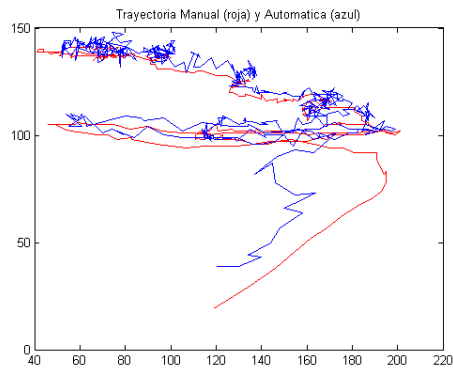


Figura 39 Grafica comparación puntos Manuales y Automáticos

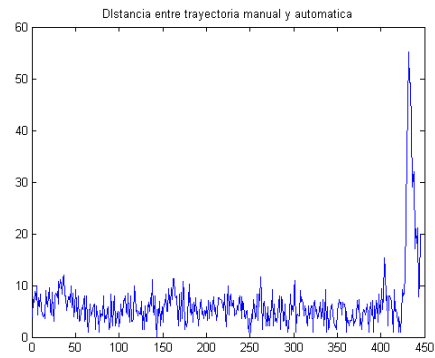


Figura 40 Distancia Euclidiana entre puntos Manuales y Automáticos

Este es el video igual al anterior, pero se obtienen peores resultados ya que el objeto al final se pierde. Esto se debe a que en ese punto que se pierde el objeto, este está aumentando mucho la velocidad.

Grafico Comparativo Algoritmos CCOEFF y SQDIFF a partir del promedio de las distancias para los mejores resultados de cada tipo de video

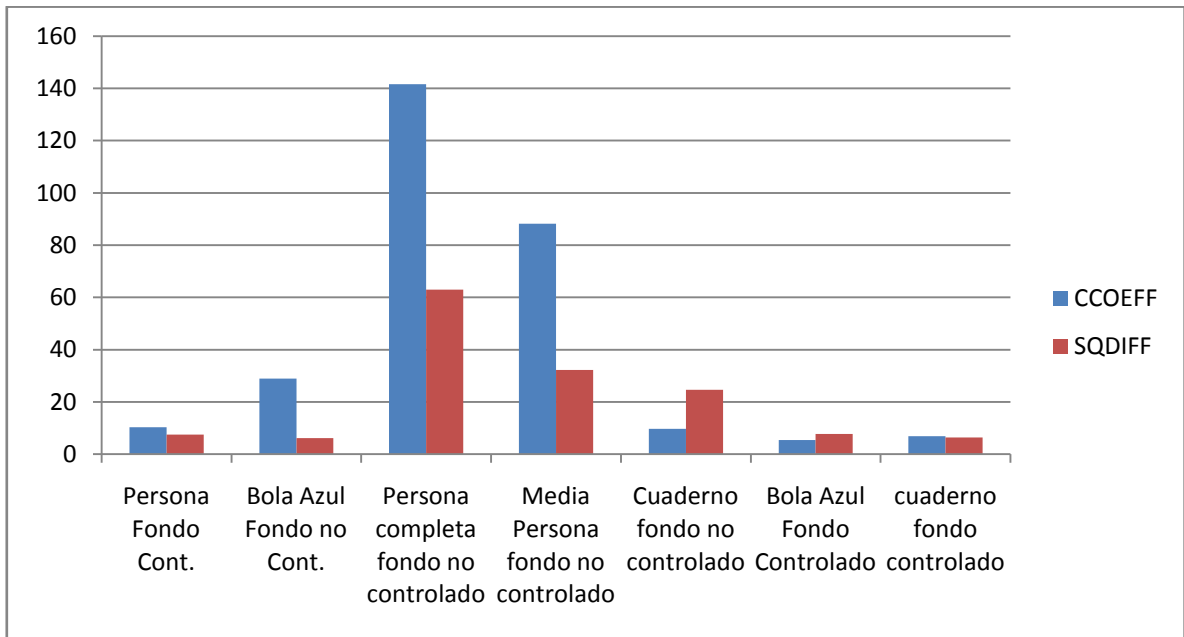
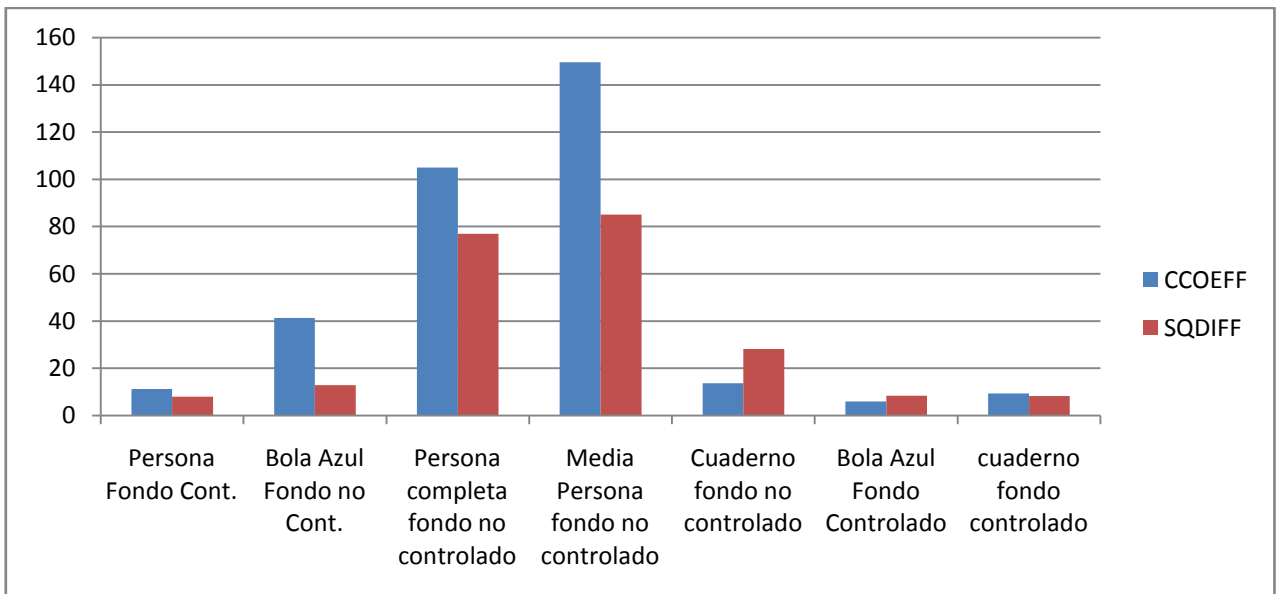


Grafico Comparativo Algoritmos CCOEFF y SQDIFF a partir de las distancias para el promedio de los resultados de cada tipo de video



En este cuadro se comparan algunos de los resultados relevantes obtenidos los demás puede verse en el ANEXO J. Claramente se observa que con el método de diferencia cuadrada se obtienen

mejores resultados. En algunos casos el método del coeficiente de correlación muestra mejores resultados, pero en general, el de diferencia de cuadrados es el que mejor funciona. Según la gráfica se puede deducir que los mejores resultados se obtienen para objetos rígidos y no flexibles, sin embargo, los videos de personas que se tomaron para realizar estas pruebas no eran buenos videos debido a la iluminación de la hora del día que se tomaron, y además unos de estos eran con personas las cuales tenían vestimenta de un color muy parecido al fondo, por esto es que no fue exitoso el seguimiento. Sin embargo para videos de personas con colores que contrasten con el fondo el algoritmo funciona correctamente.

Tabla de Tiempos de Ejecución

TIPO DE VIDEO	TAMAÑO DE LA DISTRIBUCIÓN	NÚMERO DE PARTÍCULAS	TIEMPO PROMEDIO
<i>Objeto Rígido sin Aristas Fondo Controlado</i>	20 x 20	20	85338.62749 us
<i>Objeto Rígido sin Aristas Fondo NO Controlado</i>	20 x 20	50	190119.2297 us
<i>Objeto Rígido con Aristas Fondo Controlado</i>	20 x 20	30	469340.4204 us
<i>Objeto Rígido con Aristas Fondo NO Controlado</i>	25 x 25	85	589172.1895 us
<i>Objeto Flexible Fondo Controlado</i>	20 x 20	40	1450167.504 us
<i>Objeto Flexible Fondo NO Controlado (1)</i>	30 x 30	200	1442019.627 us
<i>Objeto Flexible Fondo NO Controlado (2)</i>	25 x 25	80	1917655.438 us

Tabla 6 *Tabla de Tiempos*

6. CONCLUSIONES

Es importante tener una buena iluminación en el video, ya que cuando no la hay los videos se ven o muy claros o muy oscuros, lo cual no es recomendable debido a que no existe mucho contraste entre el objeto de interés y el fondo. En estos casos es posible que se tomen regiones del fondo como si fueran el objeto en sí. Esto conlleva a la importancia de tener un alto contraste entre el color del objeto de interés y el fondo. Si no hay contraste, por ejemplo. En el color de la vestimenta, en el seguimiento de personas. Si el contraste es bajo entre los colores de la persona y la escena por la que se desplaza, posiblemente no se tendrá un buen resultado, ya que lo más probable es que tome partes del fondo como si fueran partes de la región de interés como tal.

El tamaño de la ventana que se crea, el cual es definido por el programador, es un parámetro importante para realizar el seguimiento, ya que depende de este factor, el seguimiento puede realizarse o no exitosamente. Lo que hace esta ventana es generar una cantidad de N rectángulos dentro de ese intervalo, por lo tanto, si la ventana es más pequeña, entonces los rectángulos generados estarán más “comprimidos” entre sí, y si la ventana es grande los rectángulos estarán más “expandidos” el uno al otro. Por lo tanto, si el objeto de interés se mueve lento, una ventana pequeña será suficiente y práctica para realizar el seguimiento. En cambio si el objeto de interés tiene una velocidad rápida, entonces el algoritmo necesitará una ventana más grande donde los rectángulos puedan llegar más lejos de la posición inicial y pueda alcanzar a tomar la imagen entera del objeto de interés y este no vaya a salirse de la distribución, si el movimiento del objeto es tal que se sale de la distribución de puntos planteada, el algoritmo pierde el seguimiento del objeto

De lo anterior, la cantidad de partículas generadas son una característica importante para realizar el seguimiento. Entre mayor sea el número de partículas, se tiene una probabilidad mayor de elegir la región con mayor similitud debido a que se pueden realizar más comparaciones con más rectángulos. Sin embargo, tener muchas podría llegar a ser ineficiente, por lo tanto, debe escogerse un número de partículas con los cuales se alcance a lograr un seguimiento exitoso. Por otro lado, la cantidad de partículas se relaciona directamente con el tamaño de la ventana. Por esta razón es que a medida que se aumente el tamaño de la ventana para adecuarse a la velocidad del objeto, deben generarse a la vez mayor cantidad de partículas para cubrir una mayor área de seguimiento.

Los dos parámetros comentados en las dos conclusiones pasadas, se relacionan con la velocidad de ejecución del algoritmo. Entre menor cantidad de partículas y menor sea la ventana de distribución, el tiempo de ejecución del algoritmo será menor. Por esta razón, no se debe exagerar con la selección de estos parámetros. Lo recomendable es ajustar estos valores de tal forma que realice el seguimiento exitosamente con los valores menores posibles.

Se presenta un seguimiento satisfactorio, con menor probabilidad de error con objetos rígidos y un fondo controlado o no controlado y con las demás clases listadas de objetos con un fondo controlado.

Para trabajos futuros se recomienda partir de la característica de diferencia cuadrada, y probar con características más especializadas en el dependiendo de la clase de objeto a seguir. También se recomienda probar con distribuciones diferentes para generar la nube de partículas que dependan de la característica del movimiento del objeto.

7. BIBLIOGRAFÍA

- [1] CHENG LUO, XIONGCAI CAI, JIAN ZHANG. “Robust Object Tracking Using the Particle Filtering and Level Set Methods: A Comparative Experiment”. *School of Computer Science and Engineering, The University of New South Wales, Sydney, NSW 1056, Australia. National ICT Australia, Locked Bag 6016, NSW 1466, Australia.*
- [2] D. MANTECÓN, J.R. MARTINEZ-DE DIOS, A. OLLERO. “Técnicas Robustas de Detección y Seguimiento de Personas para Sistemas Multirrobo”. *Grupo de Robótica, Visión y Control; Camino de los Descubrimientos, sn; Sevilla.*
- [3] JANET RENO, RAYMOND C. FISHER, LAURIE ROBINSON, NOËL BRENNAN, JEREMY TRAVIS. “The Appropriate and Effective Use of Security Technologies in U.S. Schools. A guide for Schools and Law Enforcement Agencies”. *U.S. Department of Justice Office of Justice Programs. Washington, DC 20531. Mary W. Green. Sandia National Laboratories. September 1999. Chapter II.*
- [4] “Color Video Camera. Technical Manual: EVI-D70/D70P”. *2003 Sony Corporation.*
- [5] <http://opencv.willowgarage.com/wiki/> Consultado el [15-11-2009].
- [6] TINKU ACHARYA AND AJOY K. RAY. “Image Processing Principles and Applications”. *Wiley-Interscience. A John Wiley & Sons, INC., Publication.*
- [7] http://es.wikipedia.org/wiki/Modelo_de_color_RGB consultado el [07-03-2010]
- [8] BURGUEÑO RAIGAL, EVA. “Aplicación del Filtro de Partículas al seguimiento de objetos en secuencias de imágenes.” *Universidad Rey Juan Carlos. Ingeniería técnica en informática de sistemas.*
- [9] PATRINGO, JUAN JOSE. “Filtro de Partículas para seguimiento de objetos en secuencias de imágenes”. *Grupo GAVAB, Universidad Rey Juan Carlos.*
- [10] KILIAN, JOHANNES. “Simple Image Analysis By Moments” *Version 2.0.*
- [11] <http://www.flickr.com/groups/olympus-e500/discuss/72157602053629507/> [24-04-2010]
- [12] <http://docs.gimp.org/2.4/images/glossary/color-model-additive.png> [24-04-2010]

[13] <http://culturacion.com/wp-content/uploads/2008/12/cmyk-y-rgb-copy.jpg> [24-04-2010]

[14] BRADSKI. GARY & KAEHLER. ADRIAN. "Learning OpenCV" *O'Reilly. Octubre 2008*

[15] IBA, YUKITO. "Population Monte Carlo algorithms" *The Institute of Statistical Mathematics.*

8. ANEXOS

ANEXO A COLOR

A parte de la forma, textura y otras propiedades de la imagen, la información del color es una característica importante que puede utilizarse en diferentes aplicaciones de procesamiento de imágenes. Una imagen a color puede representarse como una función $C(x,y,t,\lambda)$, donde x y y son las coordenadas de la posición, λ la longitud de onda de la luz reflejada y t el tiempo en caso de imágenes dinámicas. La base de la visión del color consta de tres funciones de sensibilidad espectral las cuales son $V_R(\lambda)$, $V_G(\lambda)$ y $V_B(\lambda)$ y corresponden a los colores rojo, verde y azul. Estas tres componentes conforman el modelo RGB. Pero este no es el único modelo por el cual se puede representar una imagen, este es solamente el más común y utilizado en el procesamiento de imágenes. A parte del RGB también existen otros modelos del color por los cuales se puede representar una imagen como por ejemplo CMYK, HSV, HIS, LUV, etc. [6].

ANEXO A.1 MODELO DEL COLOR RGB

El modelo del color RGB representa los colores por medio de la composición de los colores rojo (Red), verde (Green) y azul (Blue) de acuerdo a su intensidad, es decir, una imagen consta de tres matrices en las cuales cada una representa la componente de color rojo, verde y azul respectivamente y el color de la imagen se forma a partir de la intensidad o cantidad que tenga de cada uno de estos colores en su combinación en cada píxel. La representación del modelo RGB se muestra en la Figura 41.

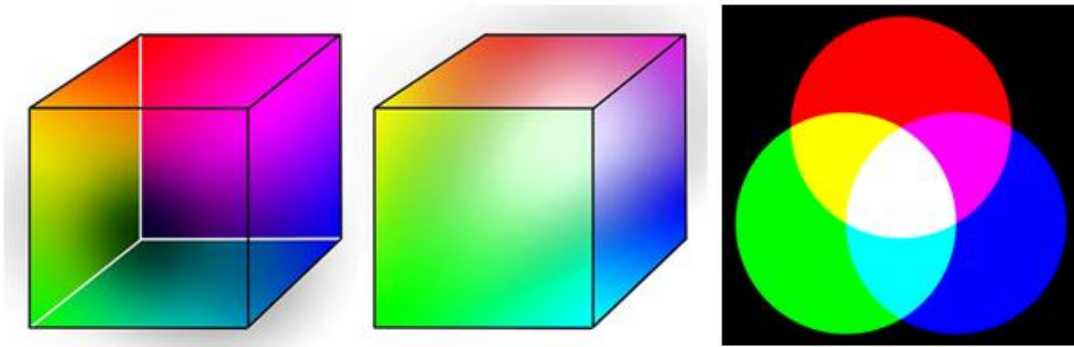


Figura 41 Representación del modelo RGB [11], [12].

Como su nombre lo dice, los colores primarios son el rojo, verde y azul. La forma de indicar la proporción de cada color que se encuentra en una mezcla, es asignar un valor a cada uno de los colores primarios, donde 0 indica que el color no se encuentra en la mezcla y este valor va aumentando a medida que aumenta su cantidad o intensidad. Para una imagen de 8 bits se tiene un intervalo de valores de 0 a 255, ya que $2^8 = 256$, pero como los valores empiezan desde 0 entonces el rango de valores llega a 255, tomando 256 valores. La manera de indicar los valores de color de un píxel es (R,G,B) . El negro es la ausencia de color, es decir, cuando los colores no aportan ninguna cantidad de color y el blanco es la intensidad máxima de color que puede presentarse, es decir, cuando los tres colores aportan con su máxima intensidad. Como se observa en la Figura 41, los colores primarios forman otros tres colores intermedios que son el amarillo, obtenido al mezclar la mayor intensidad de rojo y verde, el cian, obtenido al mezclar la mayor intensidad de verde y azul, y magenta, obtenido al mezclar la mayor intensidad de azul y rojo [7].

Típicamente, las imágenes a color, en particular las imágenes generadas por un sistema de imágenes digitales, pueden representarse mediante el modelo del color RGB.

ANEXO A.2 MODELO CMYK

Este modelo representa una imagen por medio de los colores secundarios (Cian, Magenta y Amarillo (yellow)) y el negro. Este es el modelo utilizado en las impresoras. A diferencia del RGB, la mezcla de la mayor intensidad de los tres colores secundarios forma el negro. En la Figura 42 se observa la diferencia entre los modelos RGB y CMYK.

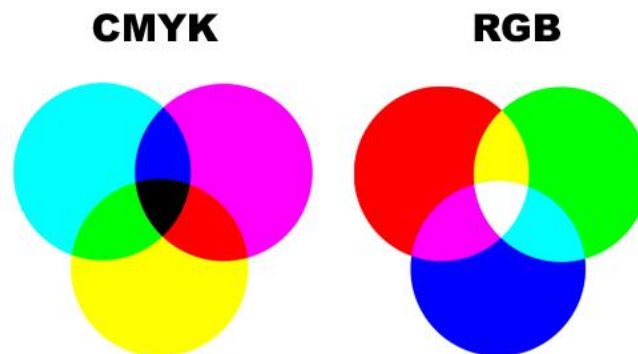


Figura 42 Modelos RGB y CMYK [13]

ANEXO A.3 MODELO HSV

Este modelo representa una imagen por medio de la información del tono del color (Hue, H), la saturación (S) y la luminancia o valor en escala de grises (V). El tono del color (H) es representado por una circunferencia donde se muestran todos los tonos de color posibles representados por el ángulo, la cual se puede observar en la Figura 43.a. La saturación (S) indica el grado de separación del gris neutral, como se muestra en la Figura 43.b. La luminancia (L) hace referencia a la cantidad de luz en escala de grises, donde 0 es el negro y 1 el blanco, como se observa en la Figura 43.c. El modelo HSV se ilustra en la Figura 43.d.

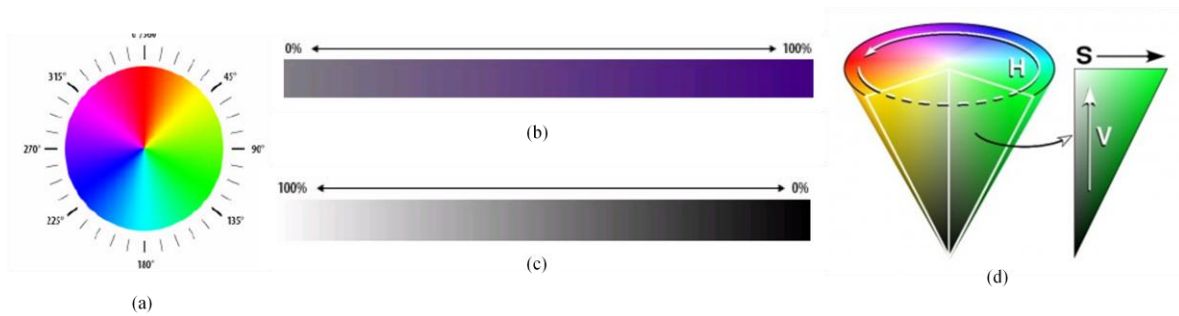


Figura 43 a) Tono del color (Hue) (b) Saturación (S) (c) Luminancia (d) Modelo HSV

ANEXO B SEGMENTACIÓN DE UNA IMÁGEN

Para realizar la interpretación y el análisis automático de una imagen es necesario desarrollar algoritmos de segmentación sobre la imagen. Cuando se implementa un algoritmo de segmentación sobre una imagen I se está dividiendo la imagen un conjunto de regiones $R_1, R_2, R_3, \dots, R_n$, donde los píxeles contenidos dentro de cada región poseen las mismas propiedades, las cuales pueden ser nivel de gris, contraste, valores espectrales o textura. La segmentación cumple las siguientes propiedades:

$$I = R_1 \cup R_2 \cup \dots \cup R_n.$$

$$R_i \cap R_j = \Phi, \forall i \neq j.$$

$$P(R_i) = P_{R_i}, \forall i$$

$$P(R_i \cup R_j) = \text{Falso}, \forall i \neq j.$$

Los algoritmos de segmentación se basan en dos propiedades básicas de los valores de niveles de grises: discontinuidad y similitud entre los píxeles. La partición de regiones de los algoritmos de discontinuidad se basa en los cambios abruptos de nivel de grises y su interés es la detección de líneas y bordes en una imagen. Un conjunto de píxeles conectados que tienen más o menos la misma intensidad homogénea conforman una región y el proceso de segmentación realiza la partición de finitas regiones las cuales conforman la escena completa. Las líneas, bordes y puntos aportan información acerca de las diferentes regiones de la imagen. Para realizar el proceso de segmentación existen diferentes técnicas [6].

Las técnicas de umbralización de los niveles de grises son métodos para realizar particiones de una imagen digital en diferentes regiones. Esta operación implica la identificación de un conjunto de umbrales óptimos, donde se elige el umbral donde se obtengan las regiones más significativas.

ANEXO C RECONOCIMIENTO DE PATRONES EN UNA IMAGEN

Después de segmentar la imagen, se realiza el reconocimiento de los objetos o regiones segmentadas en la escena. Cada objeto es un patrón y los valores de las medidas son las características del patrón. Un conjunto de objetos que poseen características casi idénticas pertenecen a determinada clase de patrones.

Se encuentran muchas características, de las cuales cada una tiene su técnica de medición. Ciertas características complejas se forman por combinaciones o distribuciones de características más simples. Algunas de las características de patrones de imágenes bidimensionales son su área, volumen, perímetro, superficie, etc. El método de medición de estas características es contando píxeles y la forma de un objeto se puede definir por sus bordes. Otras maneras de describir la forma de un patrón son los descriptores de Fourier, momentos invariantes y su eje medio. Una característica importante de un objeto puede ser su color, el cual puede ser definido en diferentes modelos de color y un objeto también se puede describir por su textura. [6]

ANEXO D ALGORITMO DE FILTRO DE PARTÍCULAS (MÉTODO DE COEFICIENTE DE CORRELACIÓN)

Ver carpeta de código anexo

ANEXO E ALGORITMO DE FILTRO DE PARTÍCULAS (MÉTODO DE DIFERENCIA CUADRADA)

Ver carpeta de código anexo

ANEXO F ALGORITMO DE PILOTO DE LA CÁMARA PTZ

Ver carpeta de código anexo

ANEXO G ALGORITMO PRUEBAS – PUNTOS A MANO

Ver carpeta de código anexo

ANEXO H ALGORITMO PRUEBAS – COEFICIENTE DE CORRELACIÓN

Ver carpeta de código anexo

ANEXO I ALGORITMO PRUEBAS – DIFERENCIA DE CUADRADOS

Ver carpeta de código anexo

ANEXO J RESULTADOS PRUEBAS DE SEGUIMIENTO

Ver carpeta Pruebas anexa al CD