

# Comparación Métodos de Tracking

Universidad Militar Nueva Granada, Vision por computador

Grupo: David Alejandro Guerrero Rodriguez Cod: 6000371

Oscar David Robles Briceño Cod: 1202105

Docente: Alexander Ceron Correa

(Dated: 2/04/2025)

Este análisis explora y compara seis algoritmos de seguimiento de objetos disponibles en OpenCV (Boosting, MIL, KCF, TLD, MOSSE, CSRT) evaluando su desempeño en términos de velocidad, robustez del seguimiento ante desafíos como oclusiones o cambios de apariencia, y precisión general.

## I. OBJETIVOS

1. Comparar: Evaluar el rendimiento relativo de los algoritmos de tracking Boosting, MIL, KCF, TLD, MOSSE y CSRT de OpenCV.
2. Implementar: Desarrollar un script en Python que permita ejecutar múltiples trackers sobre un video y recolectar datos básicos para la comparación.
3. Comprender: Entender los principios básicos, ventajas y desventajas de cada algoritmo para seleccionar el más adecuado según la aplicación.

## II. INTRODUCCION

El seguimiento de objetos (Object Tracking) es una tarea fundamental en visión por computador con aplicaciones en vigilancia, robótica, interacción humano-computadora y vehículos autónomos. Consiste en localizar un objeto específico a lo largo de una secuencia de video, dada su posición inicial. OpenCV, una biblioteca popular de visión por computador, ofrece diversas implementaciones de algoritmos de tracking, cada uno con sus propias características y compromisos entre velocidad, precisión y robustez. Este análisis se enfoca en comparar seis de estos métodos ('BOOSTING', 'MIL', 'KCF', 'TLD', 'MOSSE', 'CSRT') para entender sus diferencias prácticas y ayudar en la selección del más apropiado.

## III. MARCO TEORICO

**BOOSTING:** Basado en AdaBoost (Adaptive Boosting). Utiliza un clasificador online que combina clasificadores débiles (como Haar-like features) para distinguir el objeto del fondo.

1. Es uno de los algoritmos más antiguos. Relativamente lento. Funciona razonablemente bien con cambios lentos, pero es muy sensible a oclusiones y cambios rápidos de apariencia. Propenso a desviarse (drift).

**MIL (Multiple Instance Learning):** Aborda la ambigüedad de si una muestra de entrenamiento es real-

mente positiva. Trata el objeto como una "bolsa" de instancias (parches). Si al menos una instancia en la bolsa es positiva, la bolsa se considera positiva.

1. Más robusto que Boosting ante oclusiones parciales, ya que puede seguir funcionando si algunas partes del objeto son visibles. Generalmente más lento que Boosting.

**KCF (Kernelized Correlation Filters):** Pertenecer a la familia de filtros de correlación (Correlation Filters). Aprovecha propiedades matemáticas (matrices circulares y la Transformada Rápida de Fourier - FFT) para entrenar y detectar de forma muy eficiente. Utiliza kernels para mejorar la capacidad de discriminación.

1. Equilibrio excelente entre velocidad y precisión. Mucho más rápido que Boosting y MIL. Funciona bien en general, pero puede ser sensible a cambios de escala y deformaciones no rígidas.

**TLD (Tracking, Learning, and Detection):** Descompone el problema en tres componentes: un tracker a corto plazo (seguimiento), un detector (para re-inicializar si se pierde el objeto) y un módulo de aprendizaje (para mejorar el modelo del objeto y del fondo).

1. Diseñado para seguimiento a largo plazo. Robusto ante oclusiones y desapariciones temporales del objeto gracias al detector. Sin embargo, es computacionalmente más costoso y complejo. Puede desviarse si el detector falla repetidamente.

**MOSSE (Minimum Output Sum of Squared Error Filter):** Es uno de los filtros de correlación pioneros y más simples. Busca un filtro que minimice el error cuadrático entre la salida de la correlación real y la salida deseada (un pico gaussiano en la ubicación del objeto).

1. Extremadamente rápido (a menudo el más rápido). Muy simple de implementar. Sin embargo, es menos robusto que KCF o CSRT ante cambios de escala, iluminación y deformaciones. Buena base para entender los filtros de correlación.

**CSRT (Channel and Spatial Reliability Discriminative Correlation Filters):** Es una evolución de los filtros de correlación (similar a KCF) que incorpora conceptos de fiabilidad espacial y de canal. Utiliza un mapa de segmentación para enfocarse en las partes más fiables del

objeto (fiabilidad espacial) y pondera diferentes canales de características (como HOG, ColorNames) según su capacidad discriminativa (fiabilidad de canal).

1. Generalmente considerado uno de los trackers más precisos y robustos basados en CPU en OpenCV, especialmente bueno manejando oclusiones y deformaciones. Sin embargo, es significativamente más lento que KCF y MOSSE.

#### IV. RESULTADOS

##### Video Globo

Metodo	Velocidad	Seguimiento	Exactitud
BOOST	0.27	15 seg	23%
TLD	0.19	2 seg	2%
MIL	0.29	57 seg	93%
CSRT	0.23	74 seg	97%
MOOSE	0.72	23 seg	92%
KCF	0.43	6 seg	14%

TABLE I. Resultados video globo

##### Video Autos

Metodo	Velocidad	Seguimiento	Exactitud
BOOST	0.38	7 seg	44%
TLD	0.45	10 seg	71%
MIL	0.37	8 seg	42%
CSRT	0.43	14 seg	97%
MOOSE	0.83	4 seg	50%
KCF	0.70	3 seg	30%

TABLE II. Resultados video autopista seleccionando un auto blanco

#### V. ANALISIS RESULTADOS

Para conseguir el valor de la alteracion de la velocidad se tomó la formula

##### 1. Velocidad

$$V = \frac{V_f}{V_o}$$

Ecuacion (1)

Donde VO es duracion del video original y VF es duracion del video final Para hallar el seguimiento, se tomó el tiempo que el método siguió el objetivo durante la duracion del video, esto tambien incluye cuando el cuadro capto completamente al objeto y no parcialmente como en el video del carro que la mayoría lo tomaban, sin embargo estos tomaban parcialmente al carro y no por completo. Finalmente Para hallar la exactitud, se tuvo en

cuenta el tiempo del seguimiento del objeto y la duracion de este, quiere decir que a pesar de ser el video largo y el seguimiento dura lo mismo que el video significa que fue muy exacto, se tuvo en cuenta la formula:

##### 2. Exactitud

$$E_x = \frac{t_s}{t} * (100\%)$$

Ecuacion (2)

Donde (Ts) es el tiempo del seguimiento y (T) es la duracion del video procesado.

Segun estos datos podemos obtener que para el video del auto

1. MOSSE: Fue el mas rápido, y en el momento de que el carro seleccionado sale de pantalla es como si intentara seguirlo fuera, y se que da en el ultimo punto que pudo reconocer el carro.
2. KCF: Fue el segundo mas rápido, en el momento que el auto sale de la pantalla simplemente no reconoce mas cosas en los frames.
3. MIL: Aunque fue el segundo mas lento cuando el auto sale de la pantalla este se queda en la esquina y se mueve muy poco, como si estuviera buscando el objeto.
4. TLD: Fue el tercero mas rápido, pero cuando el carro sale de la pantalla empieza a reconocer otros objetos parecidos y cambia de objeto de forma constante.
5. CSRT: Fue el cuarto más rápido, Cuando el carro sale de la pantalla selecciona una zona muy grande, como intentando encontrarlo, aunque no selecciona ningún objeto como tal.
6. BOOSTING: Fue el mas lento y cuando el carro sale de la pantalla selecciona otro objeto cercano, luego empieza a cambiar constantemente de objeto, siempre uno cercano al que esta siguiendo en el momento.

Para el video del globo se notó que: Los resultados muestran diferencias significativas en el desempeño de los métodos de seguimiento para el video del globo.

1. CSRT destacó como el método más exacto (97%), manteniendo un seguimiento prolongado (74 seg) a pesar de su velocidad moderada (0.23). Esto lo hace ideal para aplicaciones donde la precisión es crítica, como seguimiento de objetos en movimientos complejos.
2. MIL mostró un equilibrio entre velocidad (0.29) y exactitud (93%), siendo el segundo mejor en seguimiento (57 seg). Sin embargo, su velocidad ligeramente superior a CSRT no compensa su menor

tiempo de seguimiento. MOSSE fue el más rápido (0.72) con una exactitud aceptable (92%), pero su seguimiento fue breve (23 seg), lo que sugiere que pierde el objeto fácilmente en escenas dinámicas.

3. BOOSTING y TLD tuvieron un rendimiento pobre: BOOSTING fue lento (0.27) y poco exacto (23%), mientras que TLD, aunque rápido (0.19), falló completamente en seguimiento (2 seg) y exactitud (2)
4. KCF mostró velocidad media (0.43), pero su exactitud fue baja (14%), indicando que no es confiable para objetos con cambios bruscos de trayectoria.

## VI. CONCLUSION

1. CSRT y MIL son los métodos más robustos para el seguimiento de objetos como globos, donde la precisión y persistencia son prioritarias. Aunque CSRT es más exacto, MIL ofrece un equilibrio razonable entre velocidad y desempeño, siendo útil en escenarios que requieren menor latencia.
2. BMOSSE, KCF, BOOSTING y TLD son inadecuados para este contexto. MOSSE, a pesar de su velocidad, pierde el objeto rápidamente, mientras que los demás combinaron baja exactitud con seguimiento inconsistente, demostrando ser poco fiables para objetos con movimientos no lineales o oclusiones parciales.

## VII. WEBGRAFÍA

- (1) BOOSTING (AdaBoost) Freund, Y., Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1), 119-139. <https://linkinghub.elsevier.com/retrieve/pii/S002200009791504X>
- (2) MIL (Multiple Instance Learning) Babenko, B., Yang, M. H., Belongie, S. (2011). Robust object tracking with online multiple instance learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8), 1619-1632. <https://ieeexplore.ieee.org/document/5674053>
- (3) KCF (Kernelized Correlation Filters) Henriques, J. F., Caseiro, R., Martins, P., Batista, J. (2015). High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3), 583-596. <https://ieeexplore.ieee.org/document/6870486/>
- (4) TLD (Tracking-Learning-Detection) Kalal, Z., Mikolajczyk, K., Matas, J. (2012). Tracking-learning-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(7), 1409-1422. <https://ieeexplore.ieee.org/document/6104061>
- (5) MOSSE (Minimum Output Sum of Squared Error Filter) Bolme, D. S., Beveridge, J. R., Draper, B. A., Lui, Y. M. (2010). Visual object tracking using adaptive correlation filters. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2544-2550. <https://ieeexplore.ieee.org/document/5539960/>
- (6) CSRT (Channel and Spatial Reliability Tracker) Lukezic, A., Vojir, T., Cehovin Zajc, L., Matas, J., Kristan, M. (2017). Discriminative correlation filter with channel and spatial reliability. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6309-6318. <https://ieeexplore.ieee.org/document/8099998/>