# OBJECT TRACKING UNDER NON-UNIFORM ILLUMINATION CONDITIONS

**Kenia Picos, Víctor H. Díaz-Ramírez**

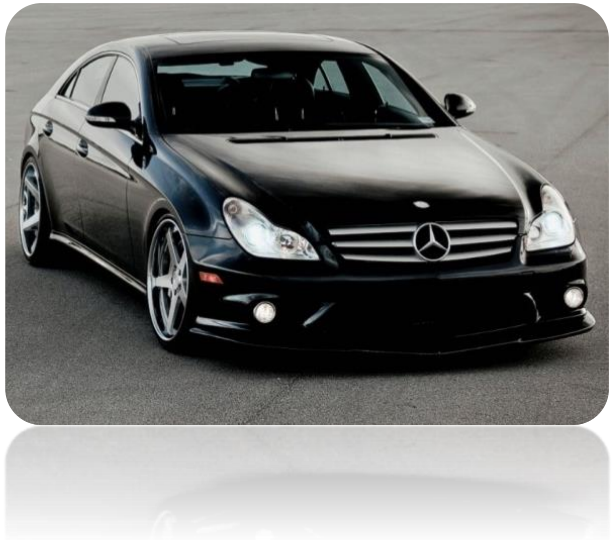
Instituto Politécnico Nacional
Center of Research and Development of Digital Technology, CITEDI
Ave. del Parque 1310, Mesa de Otay, Tijuana B.C. 22510 México

# AGENDA

- Introduction
- Pattern Recognition under non-uniform Illumination conditions
- Proposed system for Object Tracking
- Results
- Conclusions

# INTRODUCTION

Object recognition has received a notable attention due the need to construct image processing systems to improve activities such as surveillance, vehicle navigation, human-computer interaction, object tracking, among others.



# INTRODUCTION

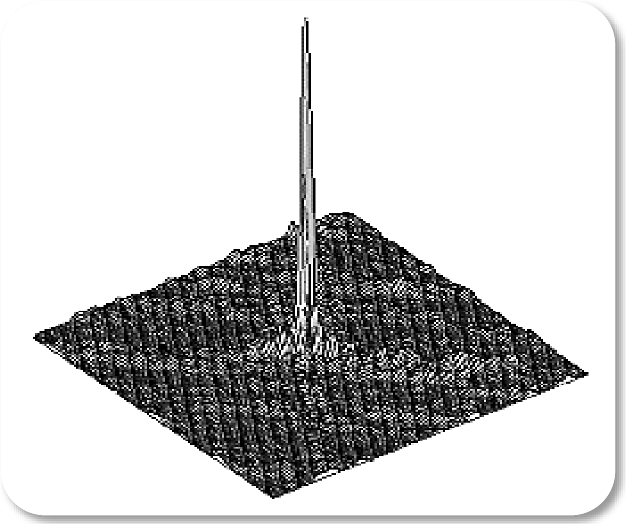
Basically, object tracking consists in solving three tasks:

Object recognition and tracking

- Object **detection** into the observed scene
- State estimation of the detected object (**location** and **orientation**)
- Object **path** estimation from input frames

# PATTERN RECOGNITION BY CORRELATION FILTERS

$\hat{x}_0, \hat{y}_0$



Estimation of target coordinates are given accurately by the coordinates of the maximum value in the output correlation plane.

However, these filters are very sensitive to scene distortions as illumination and additive noise.

It is necessary to **adapt** the impulse response of the filter for input **characteristics** of the additive noise and lighting conditions.

# ADAPTIVE PATTERN RECOGNITION

We propose an adaptive pattern recognition system able to solve the following tasks:

- Noise-tolerant Pattern Recognition
- Adaptive Light Degradation Retrieval
- Object Tracking for Movement Prediction

# PATTERN RECOGNITION BY CORRELATION FILTERS

A signal model can be expressed by

$$f(x,y) = t(x-k, y-l) + w_1(x-k, y-l)b(x,y) + n(x,y)$$

By maximization of the Signal-to-Noise ratio (SNR) criterion, the best filter for detecting the target is given by the generalized matched filter (GMF) whose frequency response is given by

$$\boxed{H^*_{GMF}(\mu,\nu) = \frac{T(\mu,\nu) + m_b W_1(\mu,\nu) + m_t W_t(\mu,\nu)}{\frac{1}{2\pi}|W_1(\mu,\nu)|^2 * N_b^0(\mu,\nu) + \frac{1}{2\pi}|W_t(\mu,\nu)|^2 * N_t^0(\mu,\nu)}}$$

Sources:

[1] V. Kober and J. Campos, "Accuracy of location measurement of a noisy target in a nonoverlapping background," *J. Opt. Soc. Am. A* **13**(8), pp. 1653–1666, 1996.
[2] B. Javidi and J. Wang, "Design of filters to detect a noisy target in nonoverlapping background noise," *J. Opt. Soc. Am. A* **11**, pp. 2604–2612, Oct 1994.

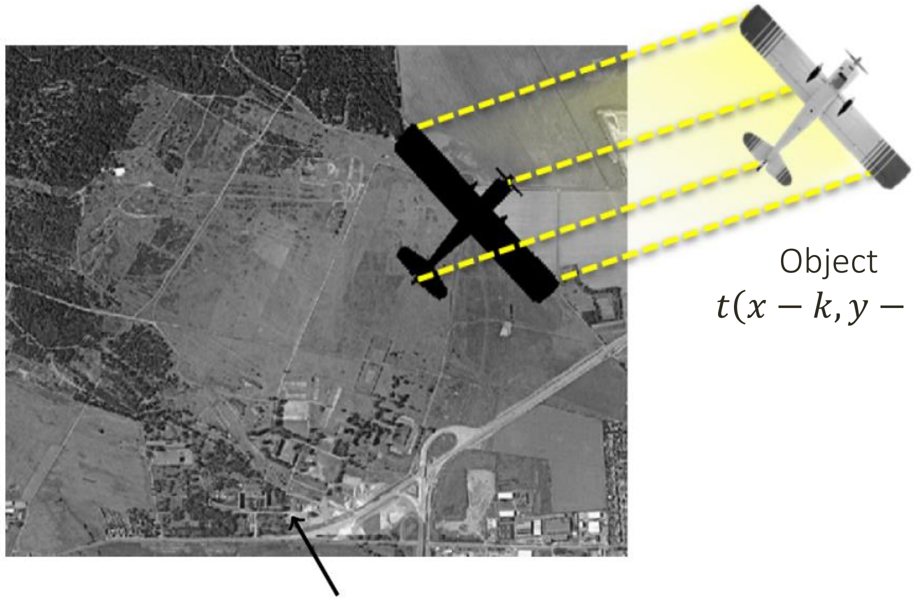
# ADDITIVE & DISJOINT NOISE-TOLERANT PATTERN RECOGNITION

Statistical parameters $\hat{\mu}_b$ y $\hat{\sigma}_b^2$ are estimated to  calculate power density function $\widehat{N}_b^0(\mu,\nu)$ required for GMF filter from an separable exponential model of the covariance function.

$$\hat{\mu}_b = \frac{\mu_f N_f - \mu_t N_w}{N_f - N_w}$$

$$\hat{\sigma}_b^2 = \frac{1}{N_b}\left(\sum_{x,y\,\in\, f}\sum f^2(x,y) - \sum_{x,y\,\in\, w}\sum t^2(x,y)\right) - \left(\frac{\mu_f N_f - \mu_t N_w}{N_f - N_w}\right)^2$$

$$\hat{S}_{b0} = (\mu,\nu) = \sum_{x,y\,\in\, f}\sum \hat{\sigma}_b^2 \rho_x^{|x|}\rho_y^{|y|} \exp(-2\pi j(\mu x + \nu y)/N_f)$$




Source:
A. K. Jain. Fundamentals of Digital Image Processing. Prentice Hall, 1988.

# LAMBERTIAN ILLUMINATION MODEL

Ligth-adaptive Pattern Recognition System

Lambertian model for 2D images with opaque surfaces

Light retrieval from statistical methods

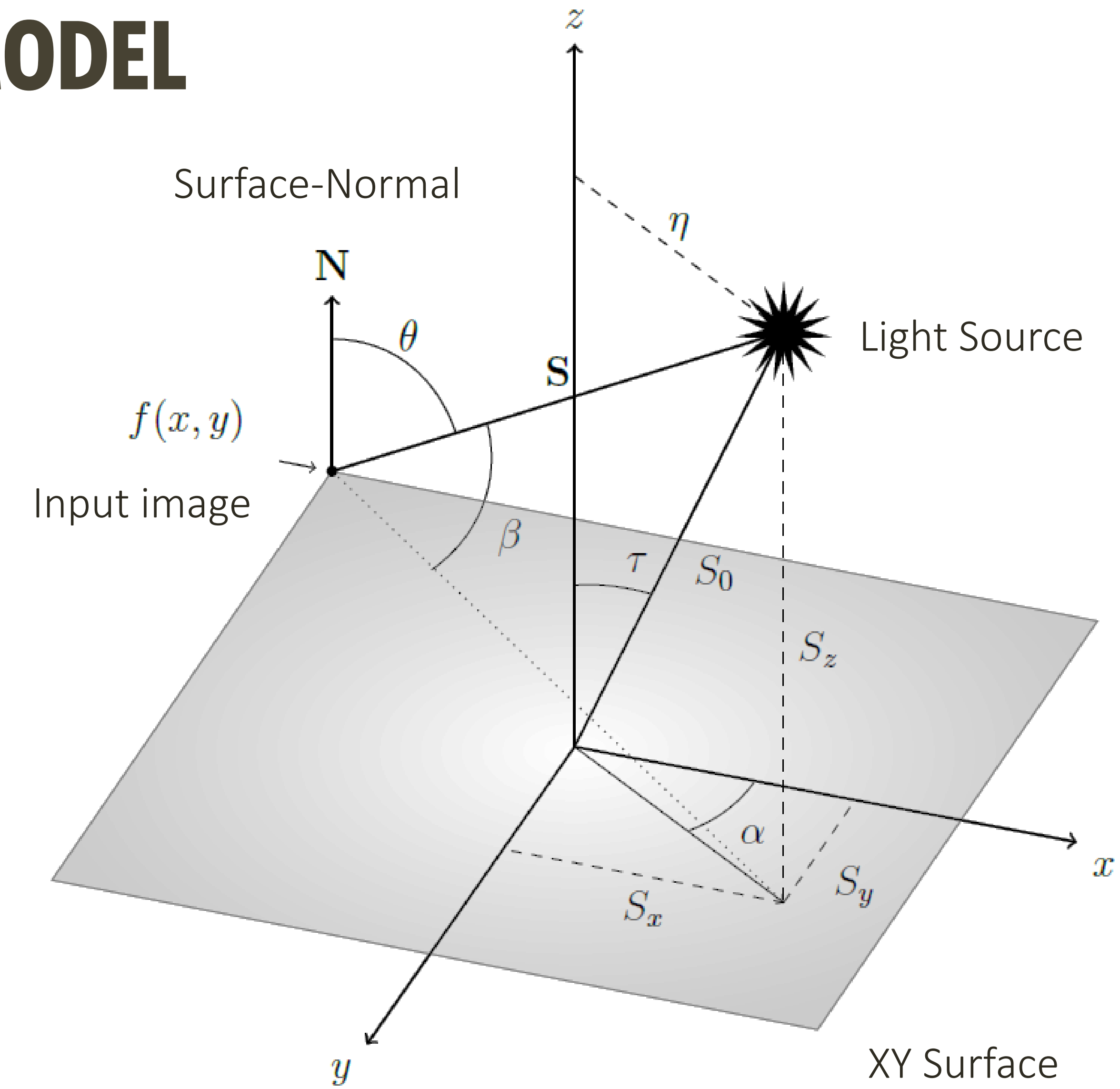
$$d(x,y) = \cos\left(\frac{\pi}{2} - \arctan\left[\frac{\rho}{\cos(\tau)}\left[(\rho\tan(\tau)\cos(\alpha) - x)^2 + (\rho\tan(\tau)\sin(\alpha) - y)^2\right]^{-1/2}\right]\right)$$

Source:

V. H. Diaz-Ramirez and V. Kober. Target recognition under nonuniform illumination conditions. Appl. Opt., 48:14081418, 2009.

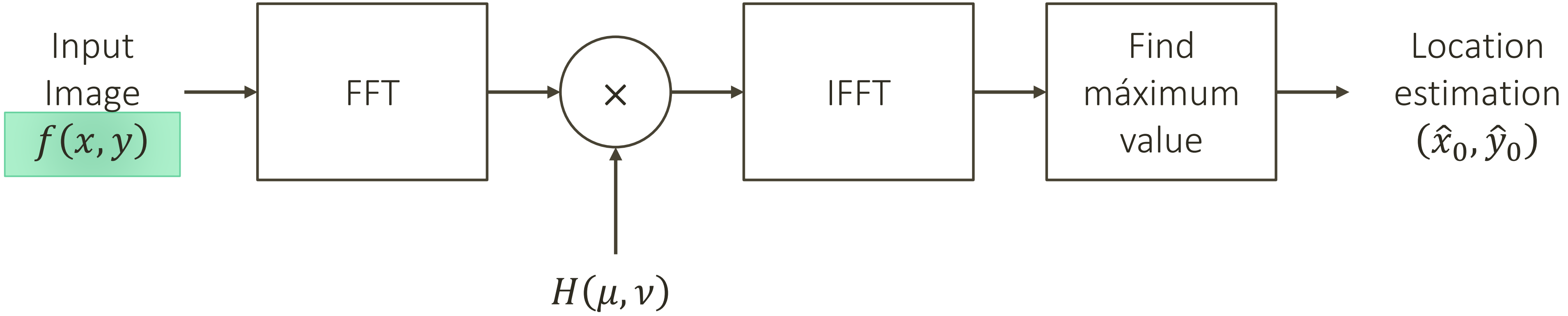
# LAMBERTIAN ILLUMINATION MODEL

- Light direction is determined for geometrical optics parameters
- Rotation angle $\alpha$ , tilt angle $\tau$, and light source distance from surface $\rho$.




Source:
V. H. Diaz-Ramirez and V. Kober. Target recognition under nonuniform illumination conditions. Appl. Opt., 48:14081418, 2009.

# ILLUMINATION MODEL




$$f(x,y) = \left[s(x-k, y-l) + b(x,y)\big(1 - w(x-k, y-l)\big)\right] d(x,y) + n(x,y)$$

Light Function

Source:
V. Kober, V. H. Díaz-Ramírez, J. González-Fraga, and J. Álvarez-Borrego. Realtime pattern recognition with adaptive correlation filters. Vision Systems – Segmentation and Pattern Recognition, pages 515536, 2007.

# NON-UNIFORM FUNCTION ESTIMATION

Statistical Image restoration

$$\hat{f}(x,y) = a_1(x,y)f(x,y) + a_0(x,y)$$

Were $a_0(x,y)$ and $a_1(x,y)$ are normalized coefficients.

Light function is unknown.

We consider the following correlation operation:

$$c(x,y) = \hat{f}(x,y) \odot h(x,y)$$

$$a_0(k,l) = \bar{t} - a_1(k,l)\,\bar{s}(k,l)$$

$$a_1(k,l) = \frac{\sum_{x,y\in w_t} t(x,y)\cdot s(k+x,l+y) - |w_t|\cdot\bar{t}\cdot\bar{s}(k,l)}{\sum_{x,y\in w_t}[s(k+x,l+y)]^2 - |w_t|\cdot[\bar{s}(k,l)]^2}$$

Were $\bar{t}$ y $\bar{s}(k,l)$ are the expected value of object and input scene in the $(k,l)$-th position, respectively inside the support region $w_t$, and $|w_t|$ is the sum of elements of $s(x,y)$.



Source:
S. Martinez-Diaz and V. Kober. Nonlinear synthetic discriminant function filters for illumination-invariant pattern recognition. Optical Engineering, 47(6):067201-19, 2008.

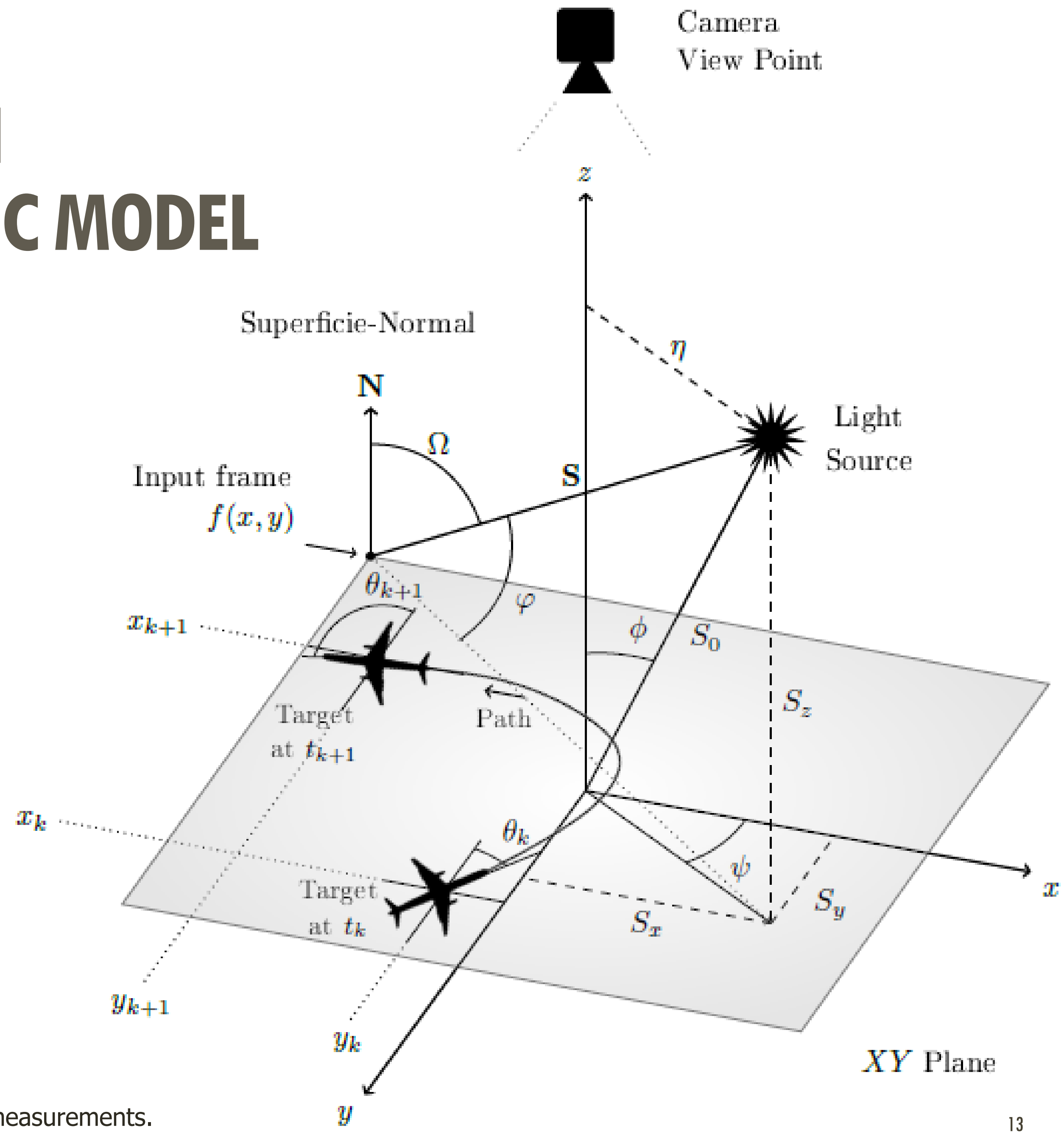
# OBJECT TRACKING AND PATH PREDICTION WITH A DYNAMIC MODEL

Object's Location and orientation prediction:

$$\hat{x}_{k+1} = x_k + \frac{\sin(\omega_k \Delta_t)}{\omega_k} \dot{x}_k - \frac{1 - \cos(\omega_k \Delta_t)}{\omega_k} \dot{y}_k$$

$$\hat{y}_{k+1} = y_k + \frac{1 - \cos(\omega_k \Delta_t)}{\omega_k} \dot{x}_k + \frac{\sin(\omega_k \Delta_t)}{\omega_k} \dot{y}_k$$

$$\phi_{k+1} = \phi_k + a_{\omega,k}$$




Source:
X. Yuan, C. Han, Z. Duan, and M. Lei. Adaptive turn rate estimation using range rate measurements. Aerospace and Electronic Systems, IEEE Transactions on, 42(4):15321541, October 2006.

# PROPOSED ALGORITHM

- **STEP 1**: Input scene
- **STEP 2**: Non-uniform illumination retrieval
- **STEP 3**: Additive and Disjoint Noise estimation
- **STEP 4**: Correlation Bank Filters Processing
- **STEP 5**: Object location and orientation estimation
- **STEP 6**: Object State prediction with a dynamic model
- **STEP 7**: Create a ROI for the new input scene

# PROPOSED ALGORITHM

Kernels in GPU

**STEP 1**

Input scene

**STEP 2**

Non-uniform illumination retrieval

**STEP 3**

Additive and Disjoint Noise estimation

**STEP 4**

Correlation Bank Filters Processing

**STEP 5**

Object location and orientation estimation

**STEP 6**

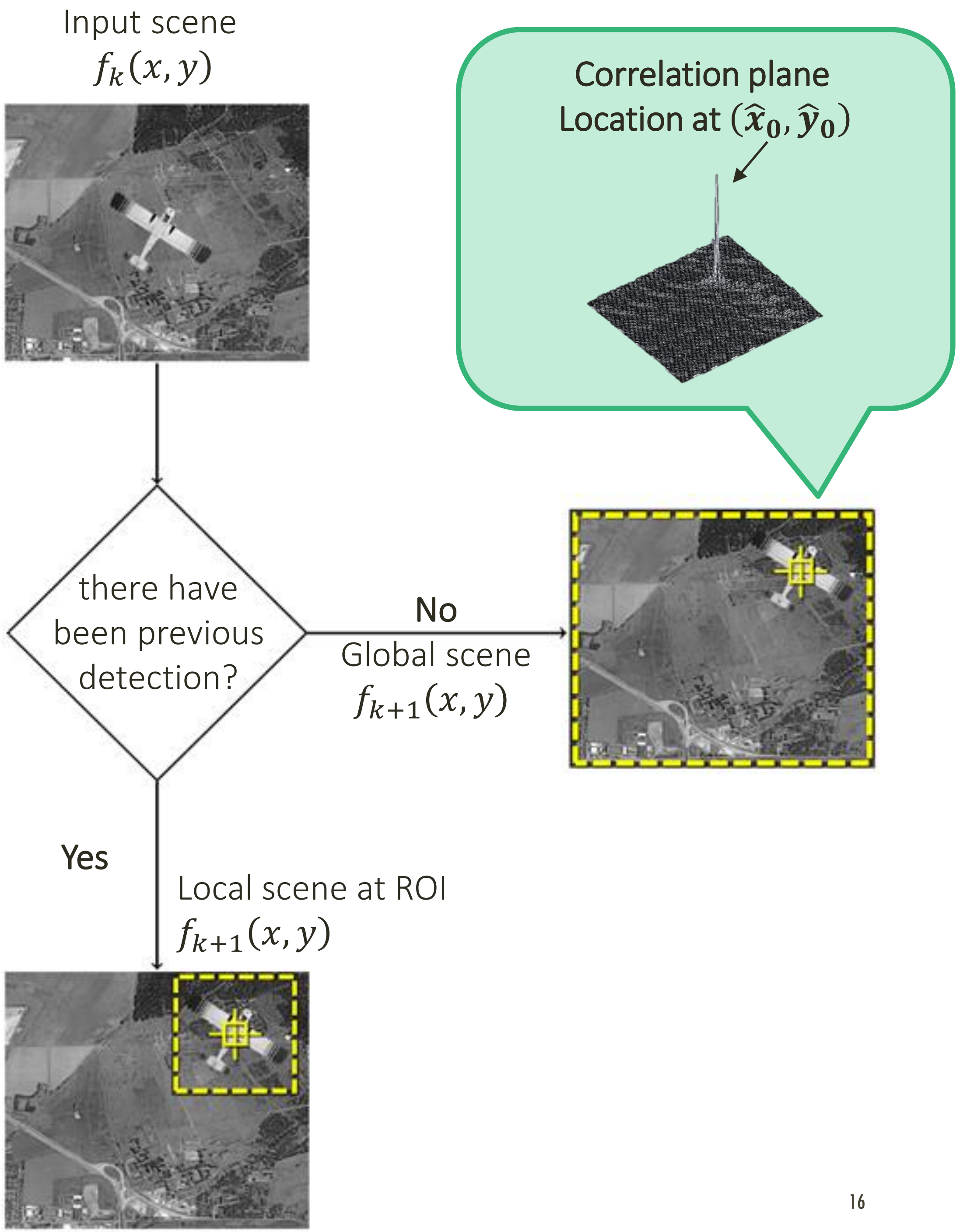
Object State prediction with a dynamic model

**STEP 7**

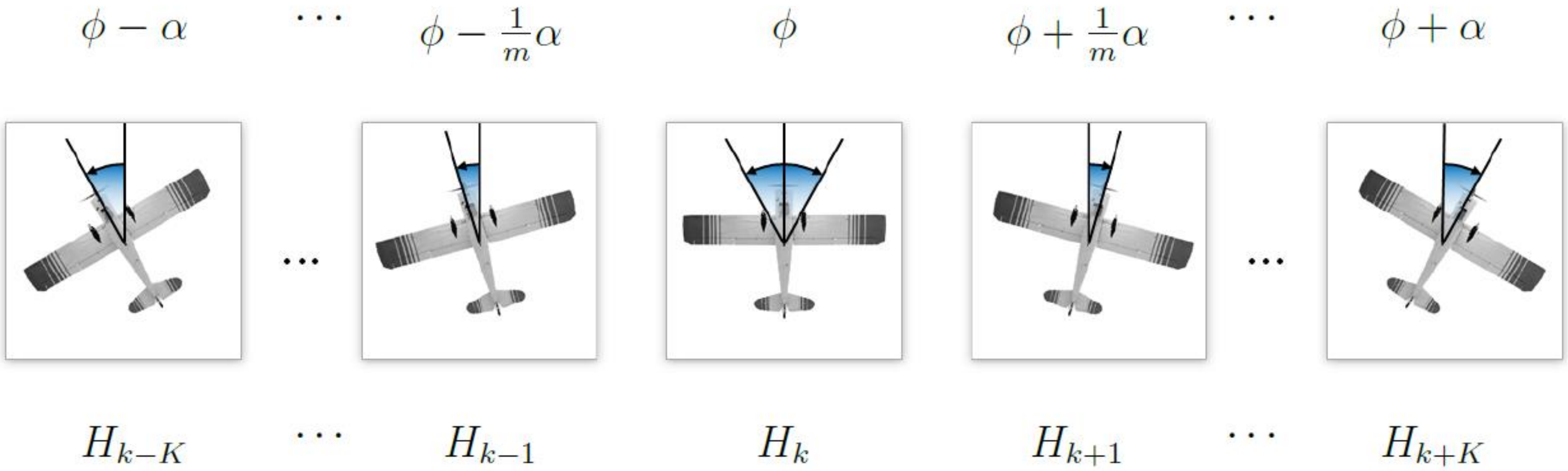
Create a ROI for the new input scene

# OBJECT LOCATION AND ORIENTATION ESTIMATION

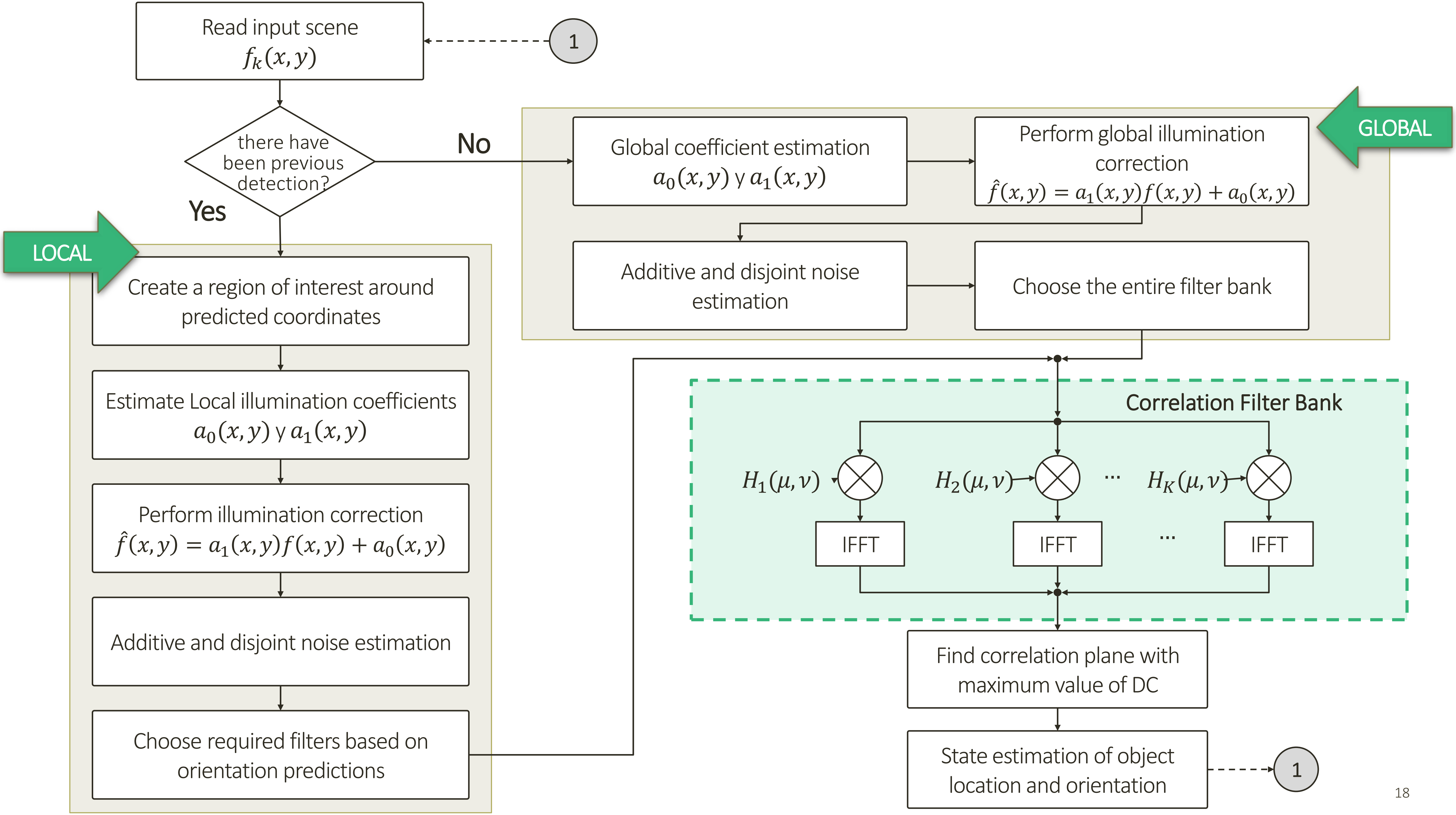
- This analysis is computed in a region of interest (ROI) given by the object's location prediction.
- This ROI can reduce the number of data used in the correlation process and the GMF filter bank.



# OBJECT'S LOCATION AND ORIENTATION ESTIMATION




- A filter bank is builded for different views of the object (rotations) .
- A system is proposed for decision taking of which filters are needed to process, based on given information of previous detections.
- It is possible to estimate the object's orientation angle, which this permits to reduce the number of operations computed in the correlation process, and the generation of GMF filters.



Read input scene
$f_k(x,y)$
1
there have been previous detection?
No
Yes
GLOBAL
Global coefficient estimation
$a_0(x,y)$ y $a_1(x,y)$
Perform global illumination correction
$\hat{f}(x,y) = a_1(x,y) f(x,y) + a_0(x,y)$
Additive and disjoint noise estimation
Choose the entire filter bank
LOCAL
Create a region of interest around predicted coordinates
Estimate Local illumination coefficients
$a_0(x,y)$ y $a_1(x,y)$
Perform illumination correction
$\hat{f}(x,y) = a_1(x,y) f(x,y) + a_0(x,y)$
Additive and disjoint noise estimation
Choose required filters based on orientation predictions
Correlation Filter Bank
$H_1(\mu,\nu)$
$H_2(\mu,\nu)$
...
$H_K(\mu,\nu)$
IFFT
IFFT
...
IFFT
Find correlation plane with maximum value of DC
State estimation of object location and orientation
1

# REAL-TIME SYSTEM IMPLEMENTATION

- Experimental platform on CPU Matlab
- GPU kernel implementation with C/C++, CUDA and ArrayFire
- GPU kernel implementation for light retrieval with C/C++, CUDA, CURAND and OpenCV

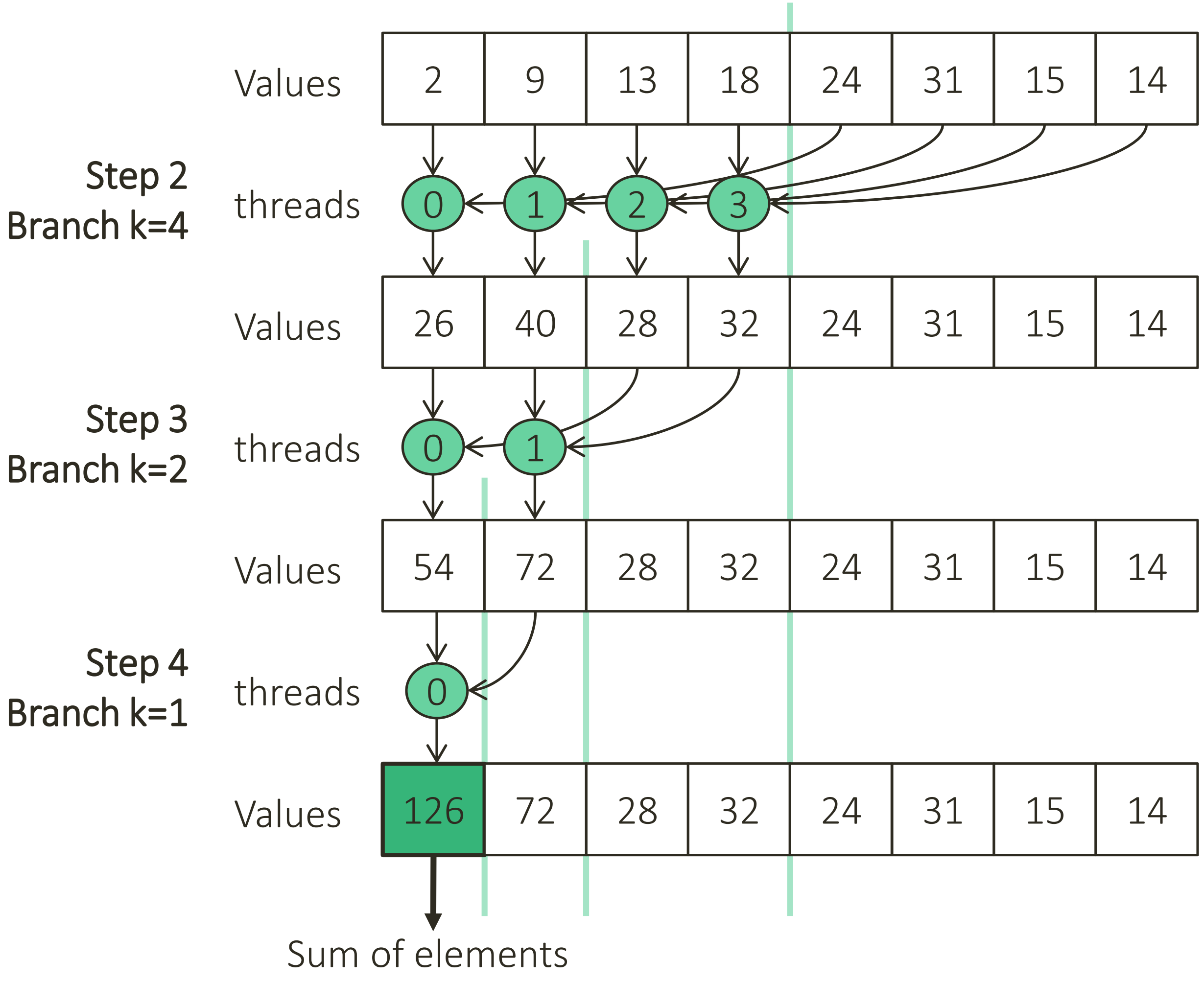
NVIDIA CUDA C/C++ ArrayFire

# LIGHT RETRIEVAL GPU IMPLEMENTATION

Two different algorithm were implemented for image restoration with **C/C++, CUDA,** CURAND API and **OpenCV.**

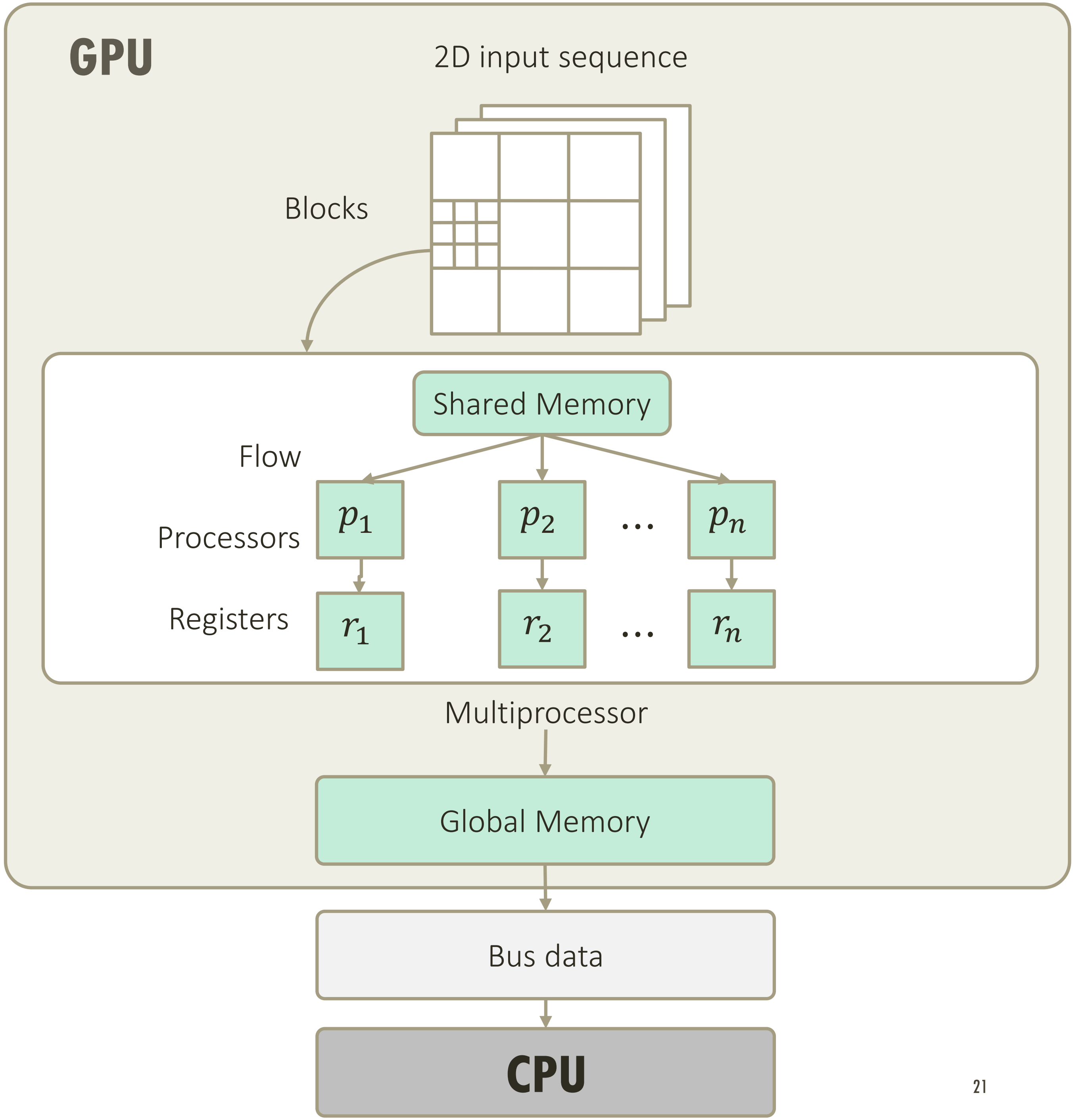
Input scenes are pixels degraded with non-uniform illumination and additive noise.

- **Kernel 1.**
  Reduction algorithm on GPU global memory
- **Kernel 2.**
  Reduction algorithm on GPU shared memory



# GPU SHARED MEMORY USAGE

Shared memory improves speedup in 4.88 x, with respect of GPU global memory usage.



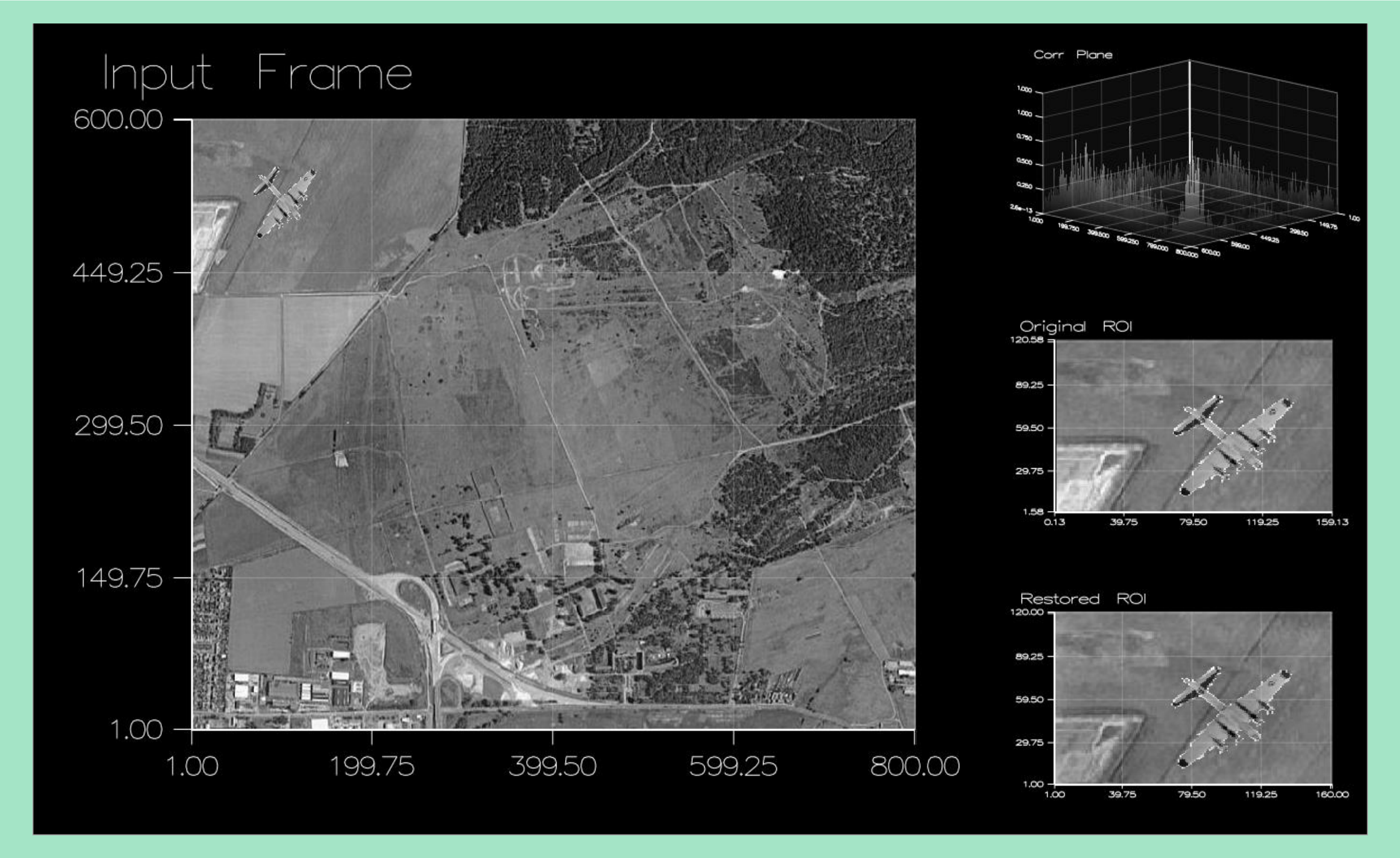
# COMPUTATIONAL PERFORMANCE

Monochromatic input images of 800×600 pixels

| | | Time | Frames per second | Speedup |
|---|---|---|---|---|
| **Kernel 1** | Global | 75.0175 s | 0.0133 fps | - - - |
| Matlab on CPU | Local | 0.3234 s | 3.0923 fps | - - - |
| **Kernel 2** | Global | 2.6426 s | 0.3787 fps | 28.3878 x |
| C/C++, CUDA, Arrayfire on GPU | Local | 0.0157 s | 63.6416 fps | 20.5977 x |

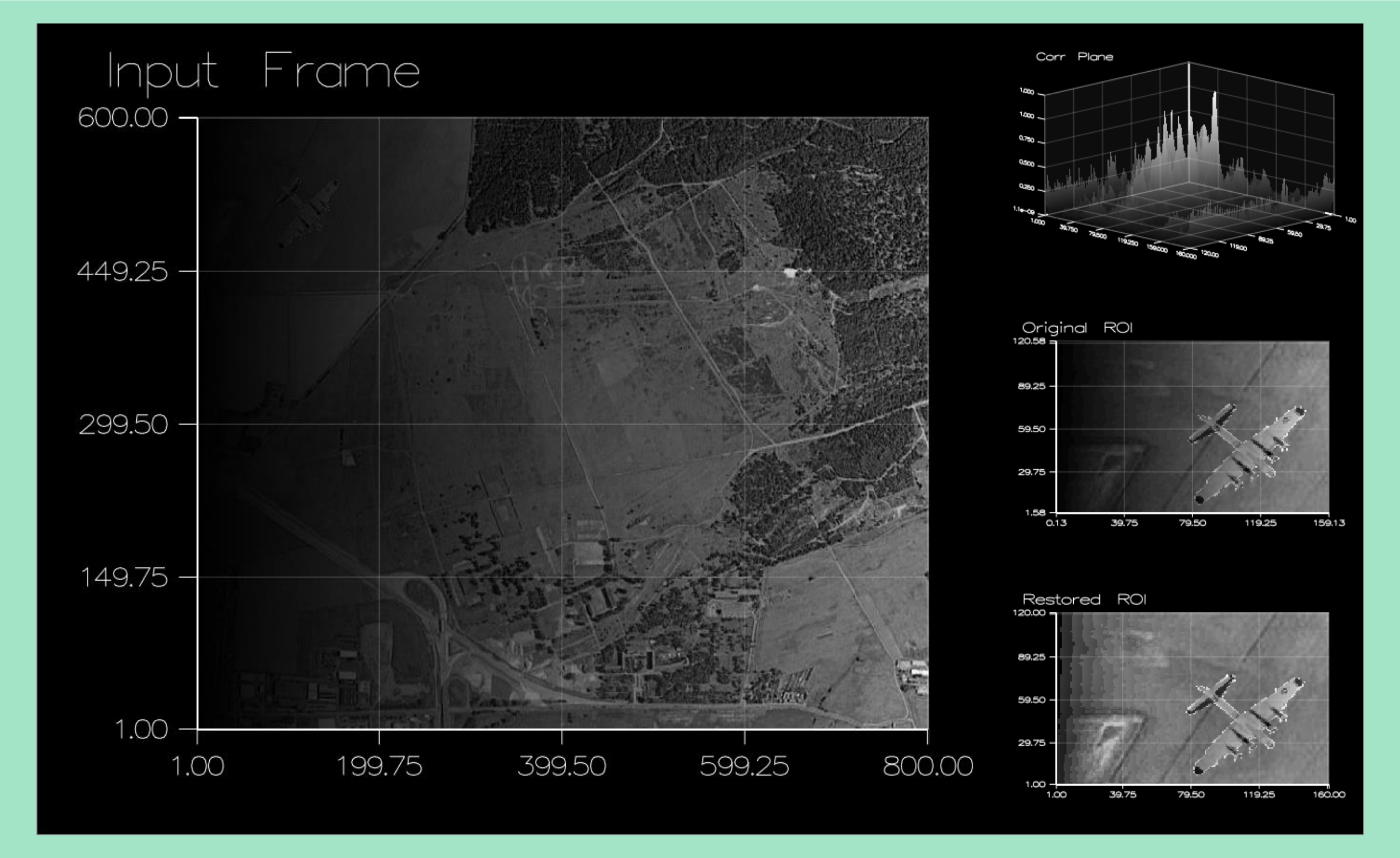
# GPU C/C++ ARRAYFIRE IMPLEMENTATION

Input scene 800x600 pixels with 45dB SNR of additive noise



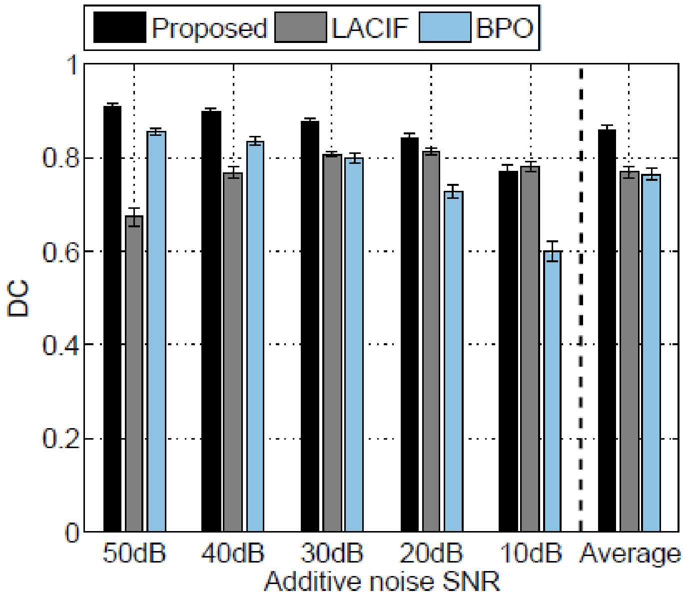
# GPU C/C++ ARRAYFIRE IMPLEMENTATION

Input scene 800x600 pixels with 45dB SNR of additive noise and non-uniform illumination degradation

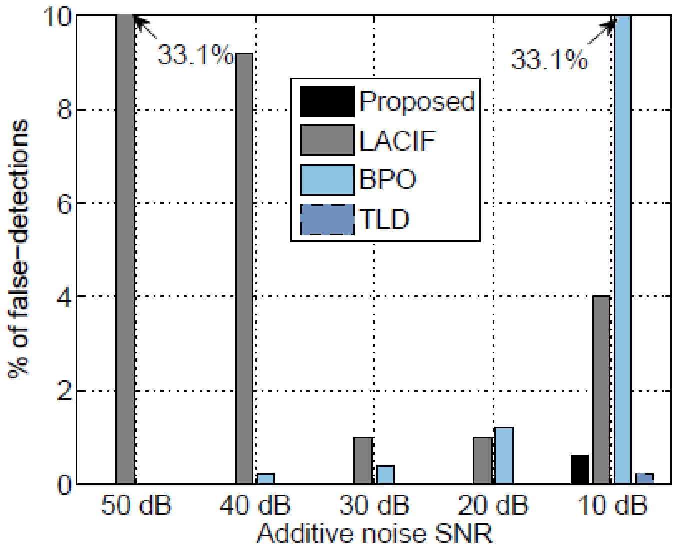


# RESULTS

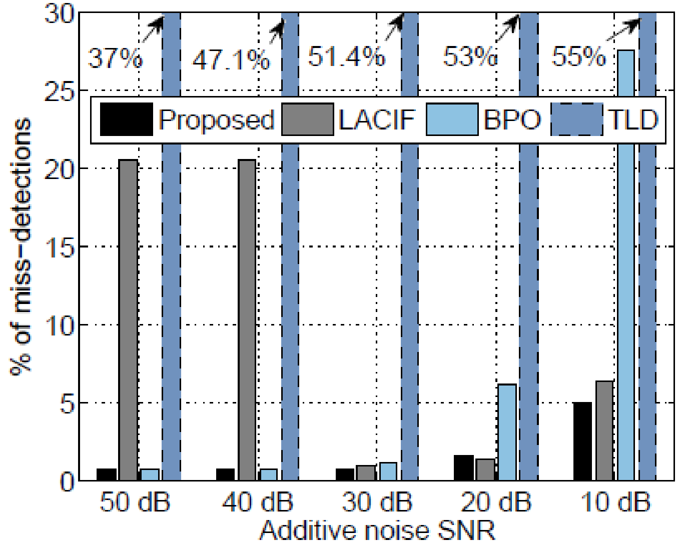
- Proposed system performance in terms of Discrimination Capability, False-Detections, and Miss-Detections.
- System comparisson with LACIF, BPO, and TLD.
- 500 frames at level of confidence=95%.




Discrimination capability



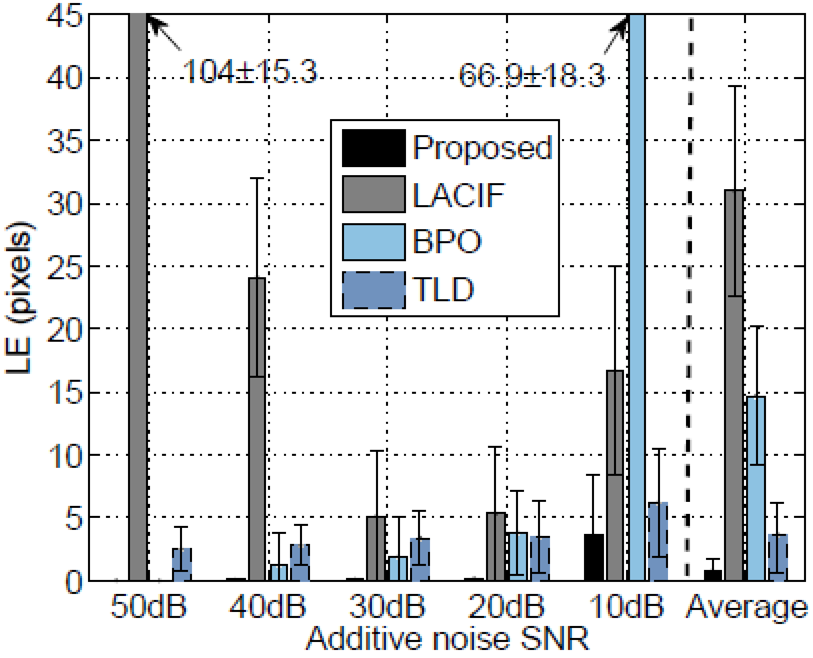

False-Detections



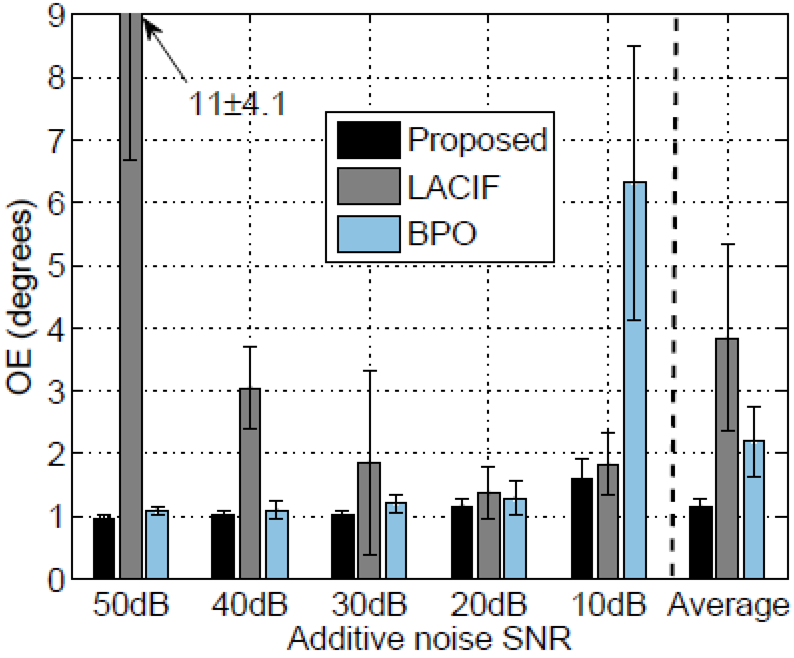

Miss-Detections

# RESULTS

- Proposed system performance in terms of Location Error, and Orientation error.
- System comparisson with LACIF, BPO, and TLD.
- 500 frames, at level of confidence=95%.




Location error




Orientation Error

# CONCLUSIONS

The system is able to correct illumination of the observed scene with a statistical restoration process.

The system generates a bank of GMF filters adaptive in statistical parameters for each frame.

The proposed system performs detection (location and orientation) in each frame with high precision.

Computer simulation results show that the proposed system performs well in terms of discrimination capability, location and orientation error in comparison with other techniques.

Image processing performance time could be improve with advanced parallel programming.

A global illumination approach will be considered in future work (reflections, different materials, 3D objects).

# THANK YOU

**Kenia Picos, Víctor H. Díaz-Ramírez**

<kpicos@citedi.mx>, <vhdiaz@citedi.mx>

Instituto Politécnico Nacional

Center of Research and Development of Digital Technology, CITEDI

Ave. del Parque 1310, Mesa de Otay, Tijuana B.C. 22510 México