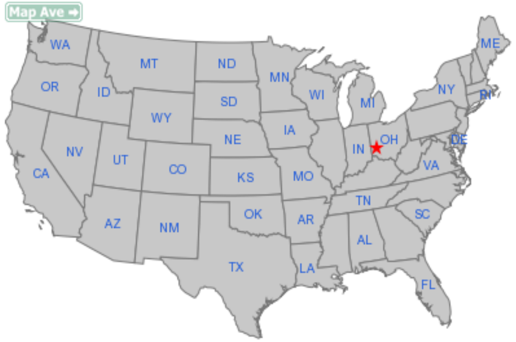
Object Detection and Recognition in Complex Environmental Conditions

Vijayan K. Asari

University of Dayton

Dayton, Ohio, USA



VISUAL 2016

13 November 2016

Overview

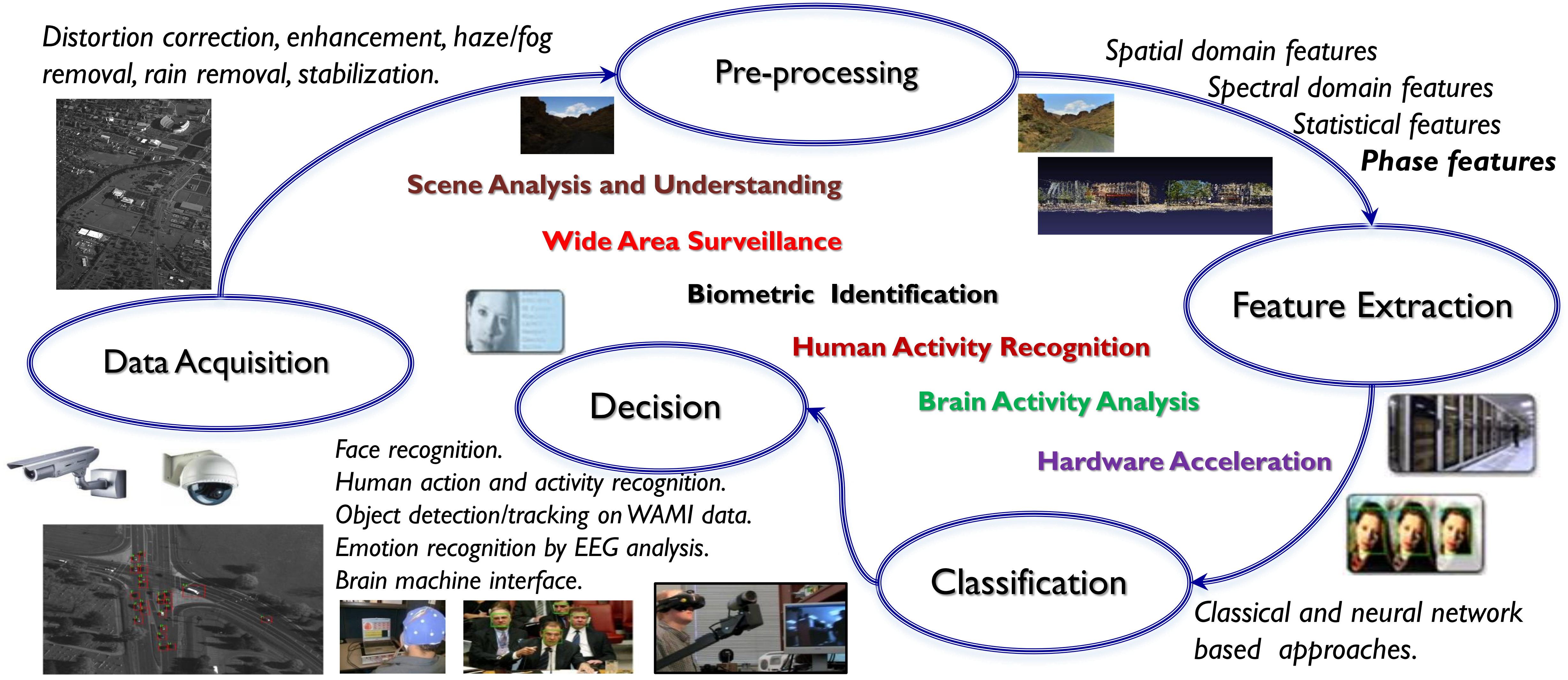
Sensor Data Acquisition → Sensor Data Exploitation → Decision and Feedback

Canon Cameras

Hyperspectral Camera

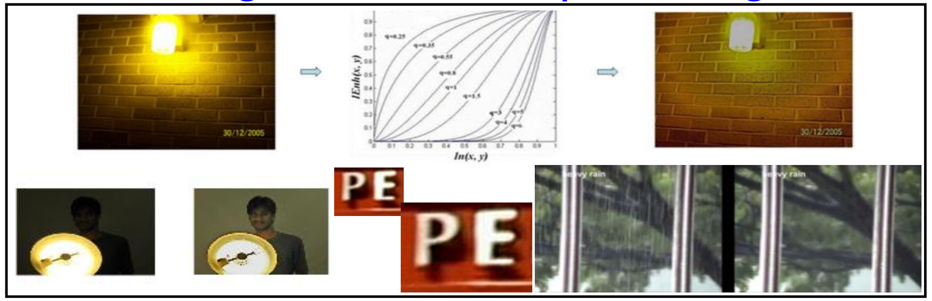
Long-range Cameras

Object Detection and Recognition: Processing Pipeline



Focus Areas

Image and Video Preprocessing

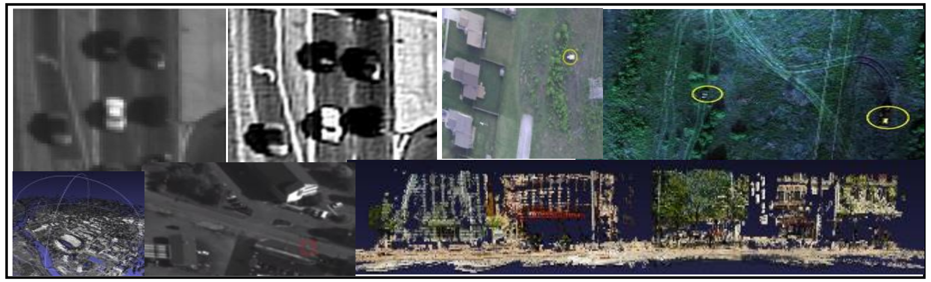


Enhancement
Super-resolution
Haze removal
Rain removal
Stabilization

Vision-Guided Robotics

Robotic navigation
Path planning
Object following
Behavior analysis
Threat analysis

Wide Area Surveillance

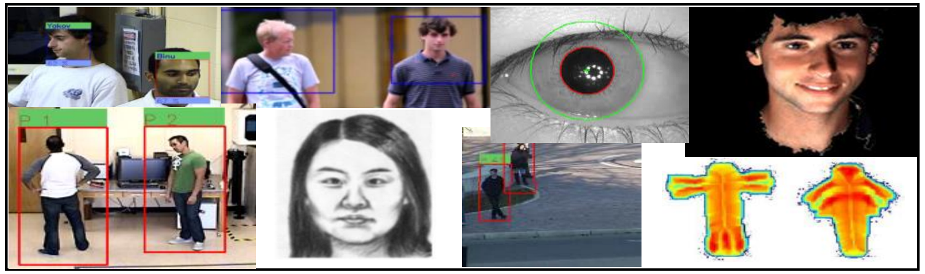


Object detection
Object recognition
Object tracking
3D reconstruction
Change detection

Perception Beyond Visible Spectrum

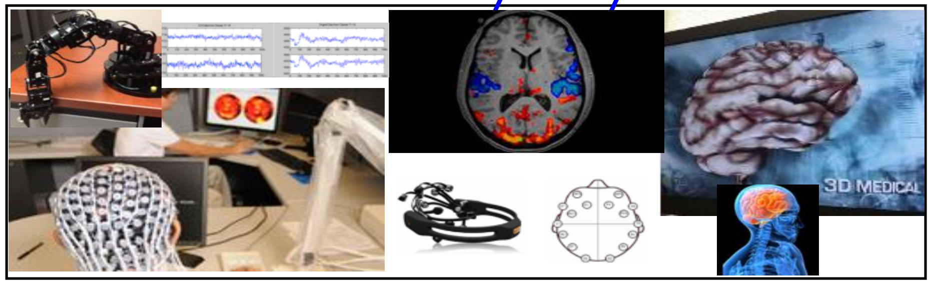
LiDAR data analysis
Hyperspectral data
IR/thermal data
Satellite imagery
EEG data analysis

Biometrics



Face recognition
Human action and activity recognition
Expression analysis
Emotion recognition

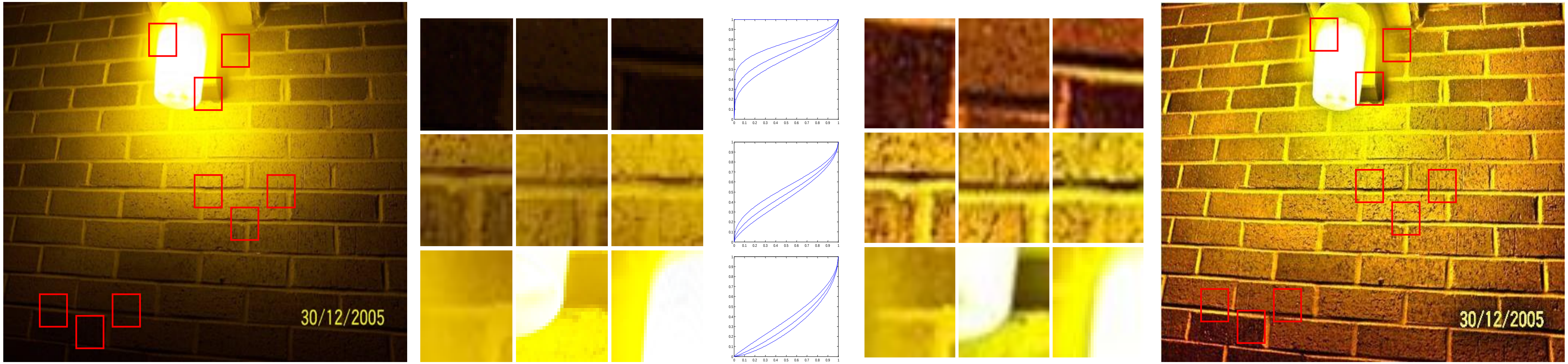
Brain Activity Analysis

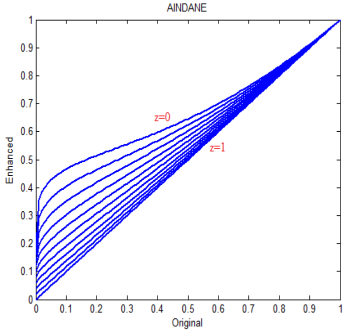


Emotion recognition
Brain machine interface
Source localization
Neurofeedback

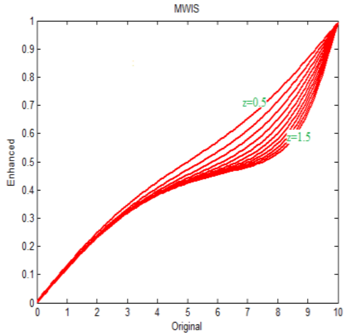
Enhancement of Low Lighting and Over Exposed Images

Underexposed, dark, dark and bright (shadows), bright, overexposed regions

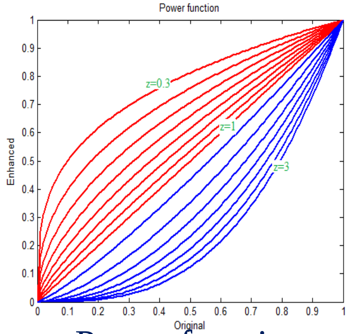




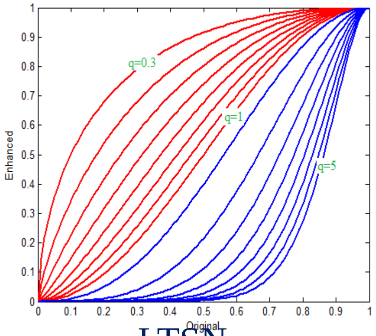
AINDANE



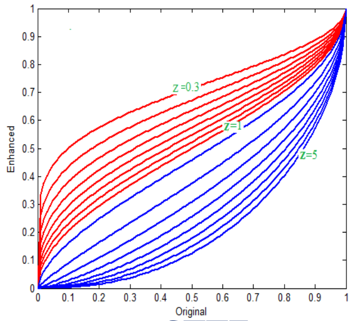
MWIS



Power function



LTSN



STTF

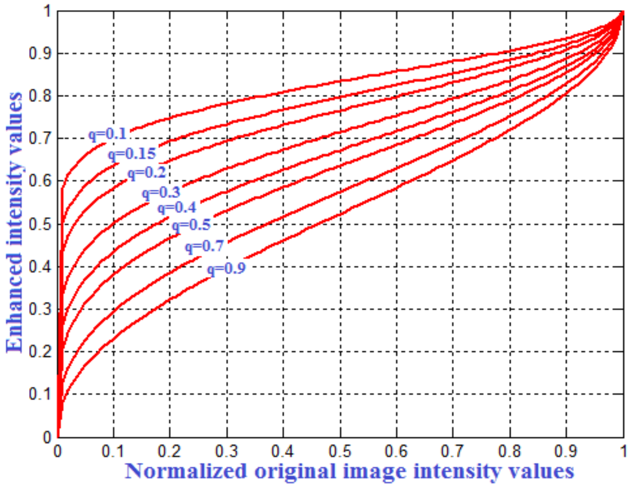
Dynamic Range Compression

Intensity computation (NTSC)

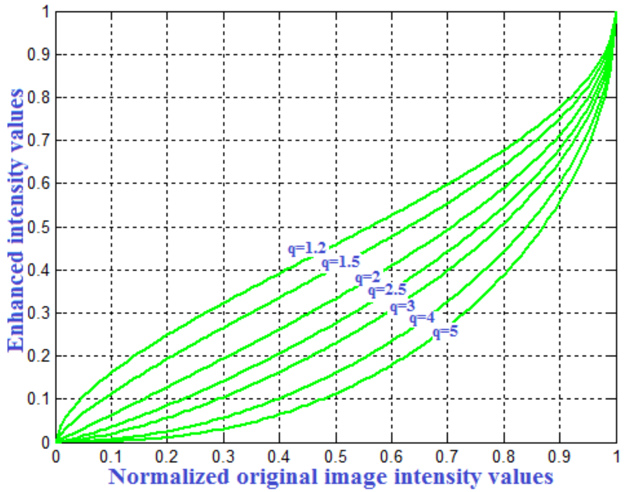
$$I(x,y) = 0.2989 \times I_{Rh}(x,y) + 0.5867 \times I_{Gh}(x,y) + 0.114 \times I_{Bh}(x,y)$$

Nonlinear function

$$I_{enh}(x,y) = (2/\pi) ArcSin(I_n(x,y)^{q/2})$$



Dark pixels



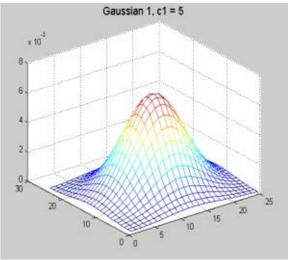
Bright pixels

Adaptive Estimation of Control Parameter

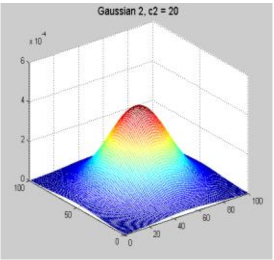
$q < 1$ Provide various nonlinear curves if the pixels are dark.

$q = 1$ Provides a curve if the pixel has sufficient intensity.

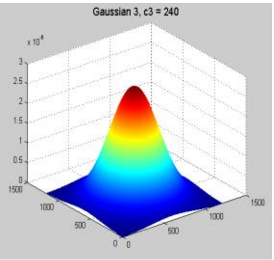
$q > 1$ Provide various nonlinear curves if the pixels are bright.



$w_1 = 5$



$w_2 = 20$

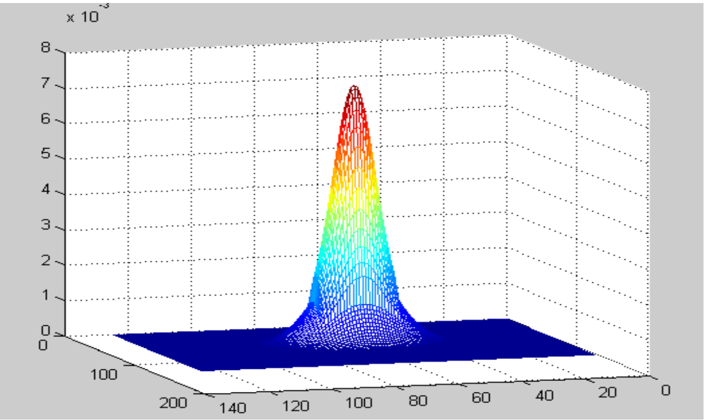


$w_3 = 240$

Depending on the mean value of its neighborhood

$$I_{M_i}(x,y)=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I(m,n)G_i(m+x,n+y)$$

$$G_i(x,y)=K.e^{\left(\frac{-(x^2+y^2)}{w_i^2}\right)}$$



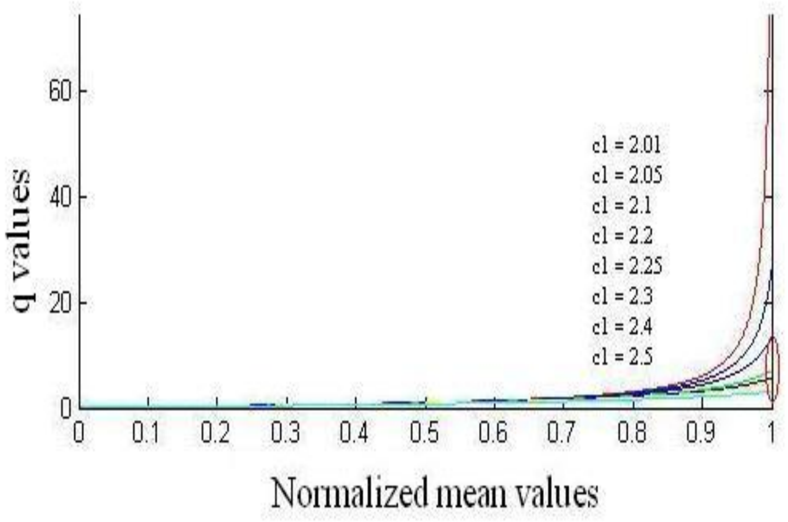
Multi-level Gaussian function

Window size depends on the resolution and object size in an image.

Adaptive Estimation of Control Parameter

Criteria for estimation of *q*

$$q = \begin{Bmatrix} <1, if I_{M_n} < 0.5 \\ =1, if I_{M_n} = 0.5 \\ >1, if I_{M_n} > 0.5 \end{Bmatrix}$$



The function for the q value can be designed as

$$q = Tan\left(I_{M_n}(x,y) * (\pi / c_1)\right) + c_2$$

c_1 and c_2 are empirically determined. $c_1 = 2.25$ $c_2 = 0.0085$

For q values which are closer to 0 the noise in the extreme dark regions will also be enhanced.

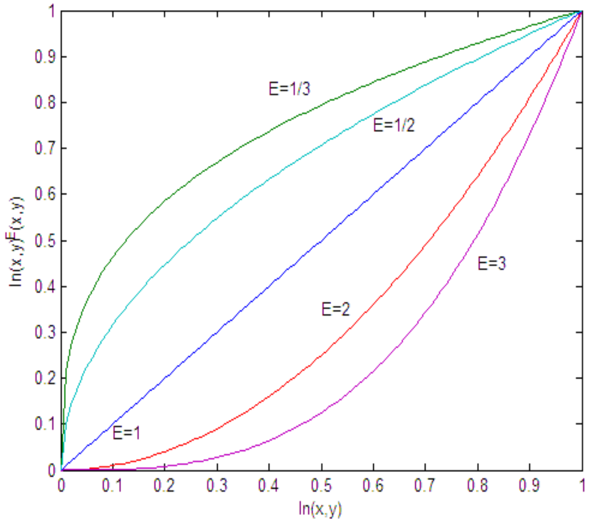
Hence, the q values corresponding to the mean value below 0.2 is considered as extreme dark regions and q for those pixels can be calculated as

$$q = \log\left(\sqrt{2 I_{M_n}(x,y) + 2}\right);$$

Nonlinear Enhancement Module

Contrast Enhancement

$$S(x,y) = 255\, x I_{enh}(x,y)^{E(x,y)} \qquad E(x,y) = \left[\frac{I_{conv}(x,y)}{I(x,y)}\right]$$



Color restoration

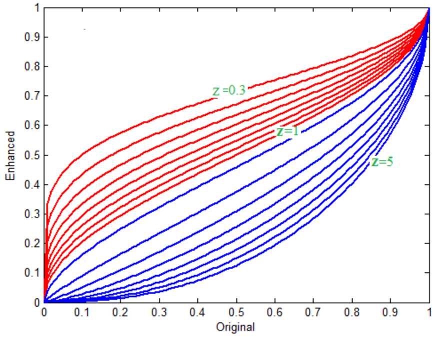
$$I_{enh,i} = I_i(x,y)\left(I_{enh}(x,y) \Big/ (I_n(x,y) \right)$$

where i represents red, green, blue spectral band

Enhancement of Low Lighting and Over Exposed Images


Input image

Enhanced image




Enhancement of Hazy/Foggy Images

Weather Degraded Image: Poor contrast, distorted color

Hazy image

Weather degraded image

Hazy image

Estimation of approximate thickness of haze in the scene and enhancement using a single nonlinear function.

An adaptive estimation of control parameter from its neighborhood information.

Enhancement of Hazy/Foggy Images

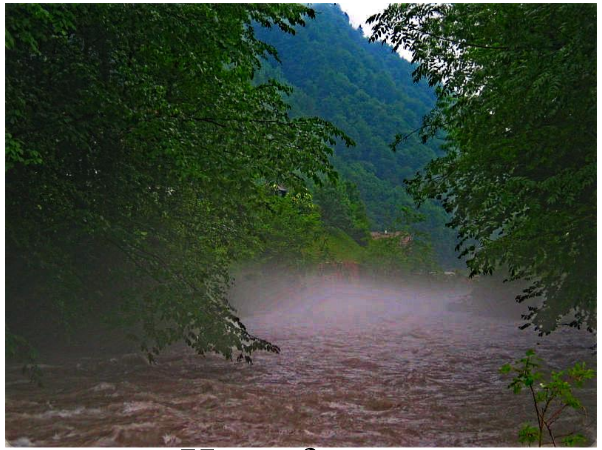

Original


Transmission map


Haze-free


Original


Transmission map

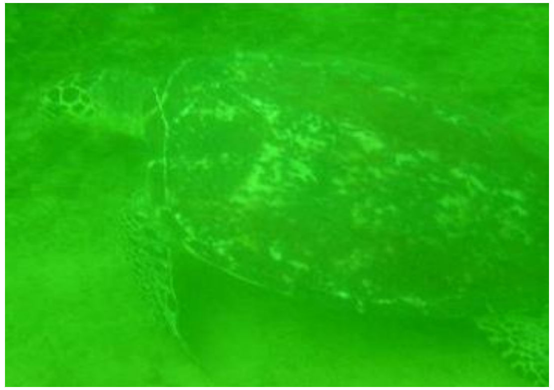

Haze-free

Enhancement of Hazy/Foggy Images





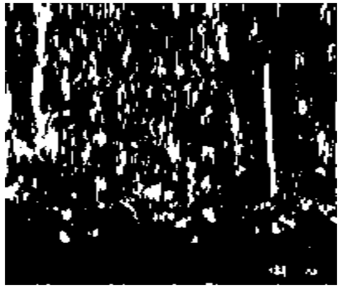

Enhancement of Hazy/Foggy Images

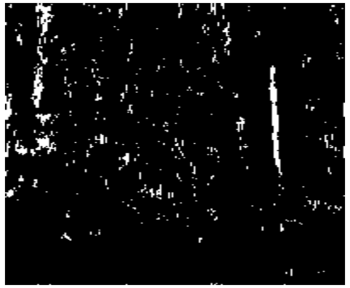


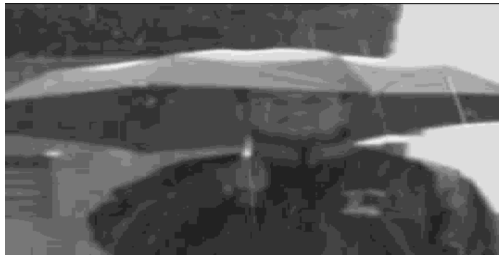
Original Images

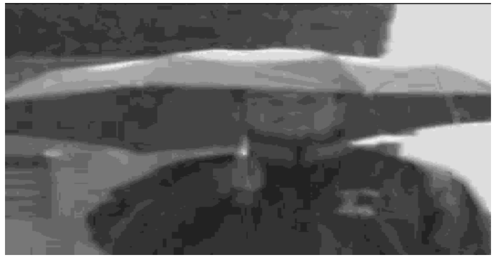
Enhanced Images

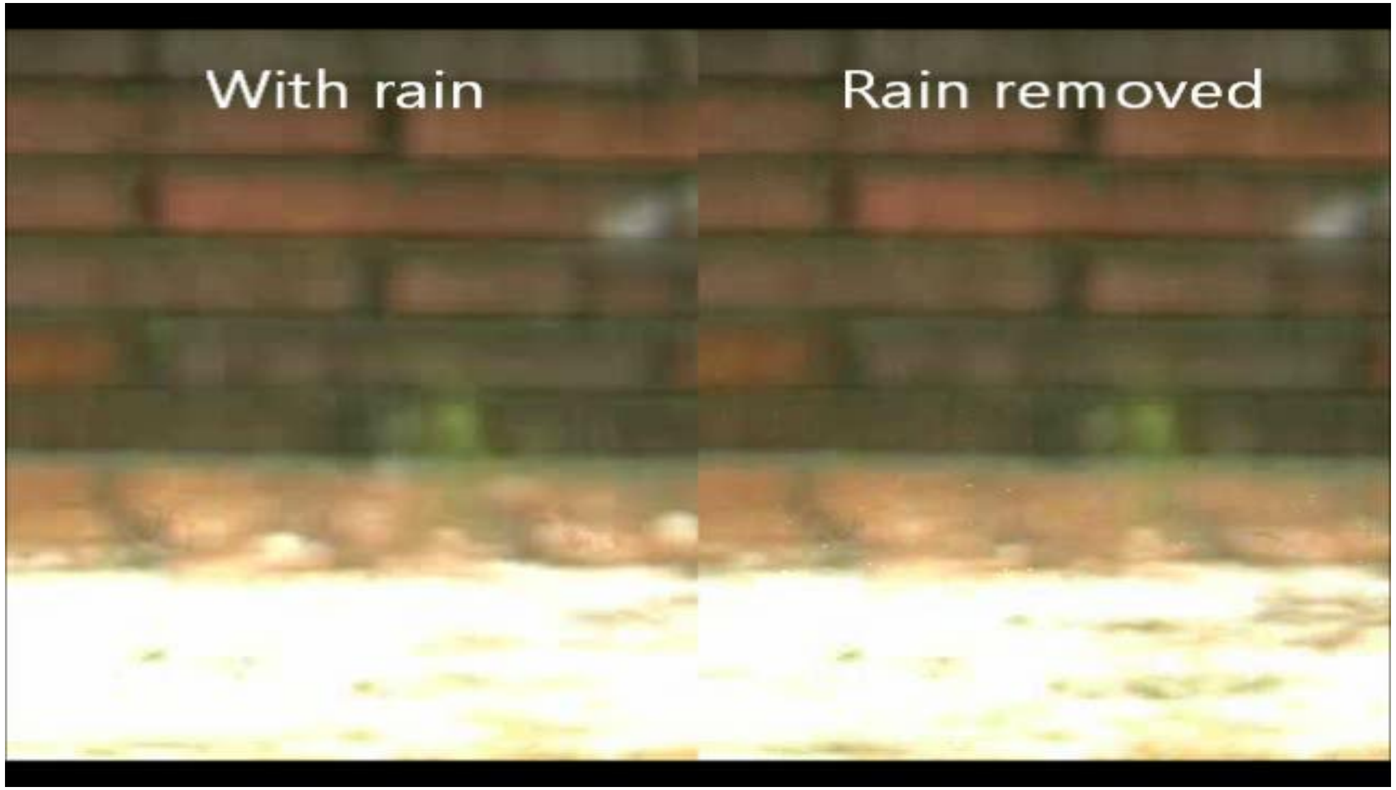
Scene Visibility Improvement: Rain Removal











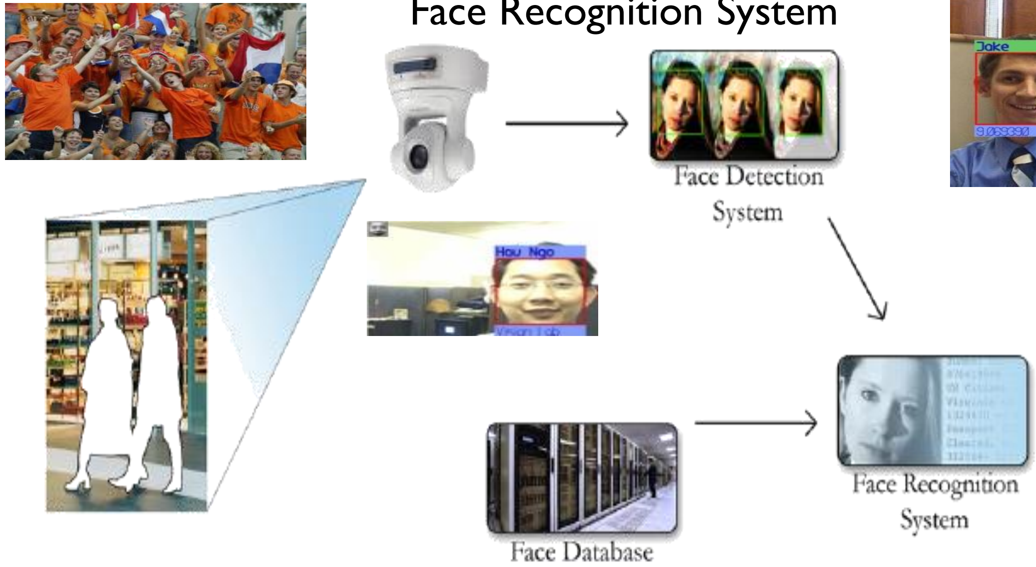
Biometric Data Analysis for Human Identification

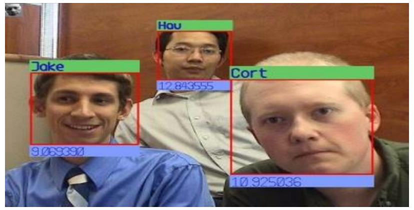
Face Detection

Action Recognition

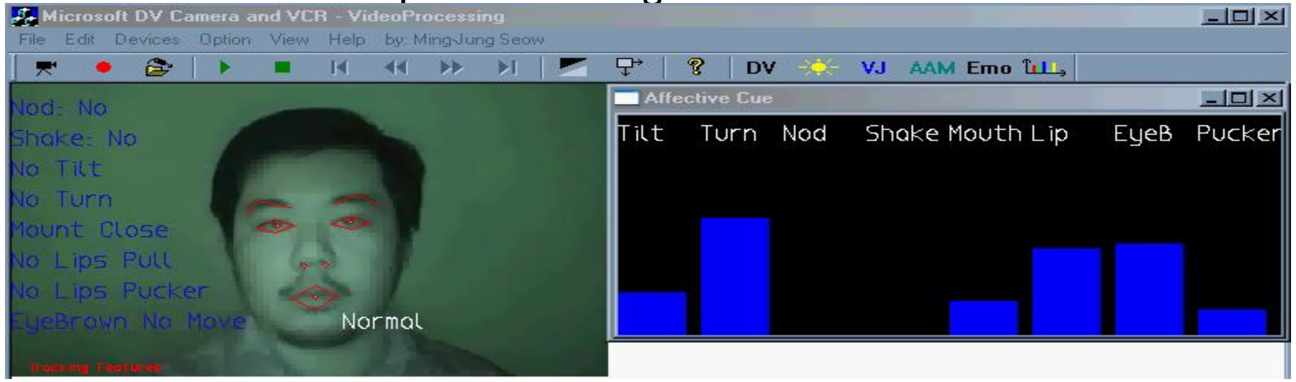
Iris Recognition

Face Recognition System

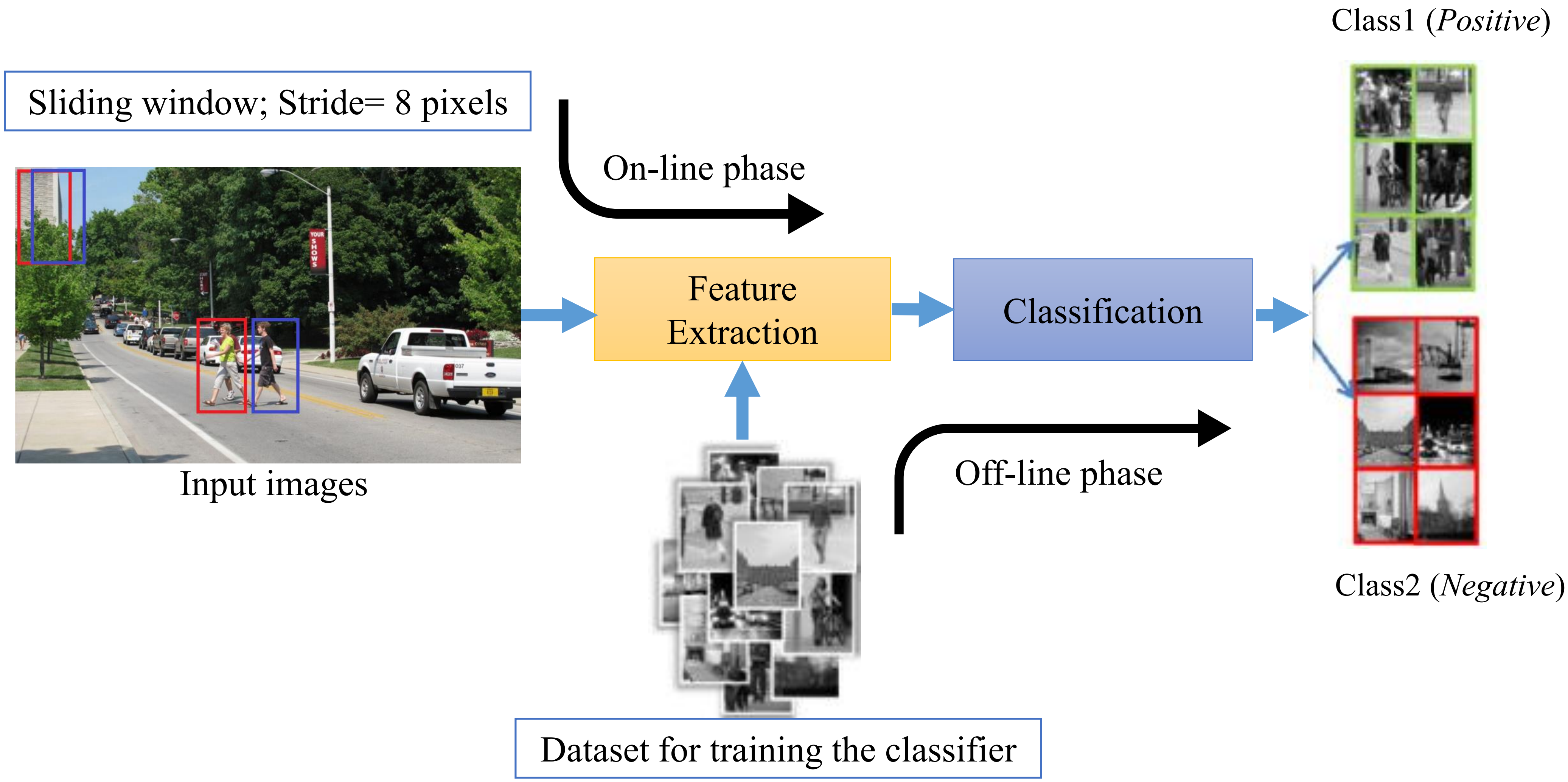




Expression Recognition

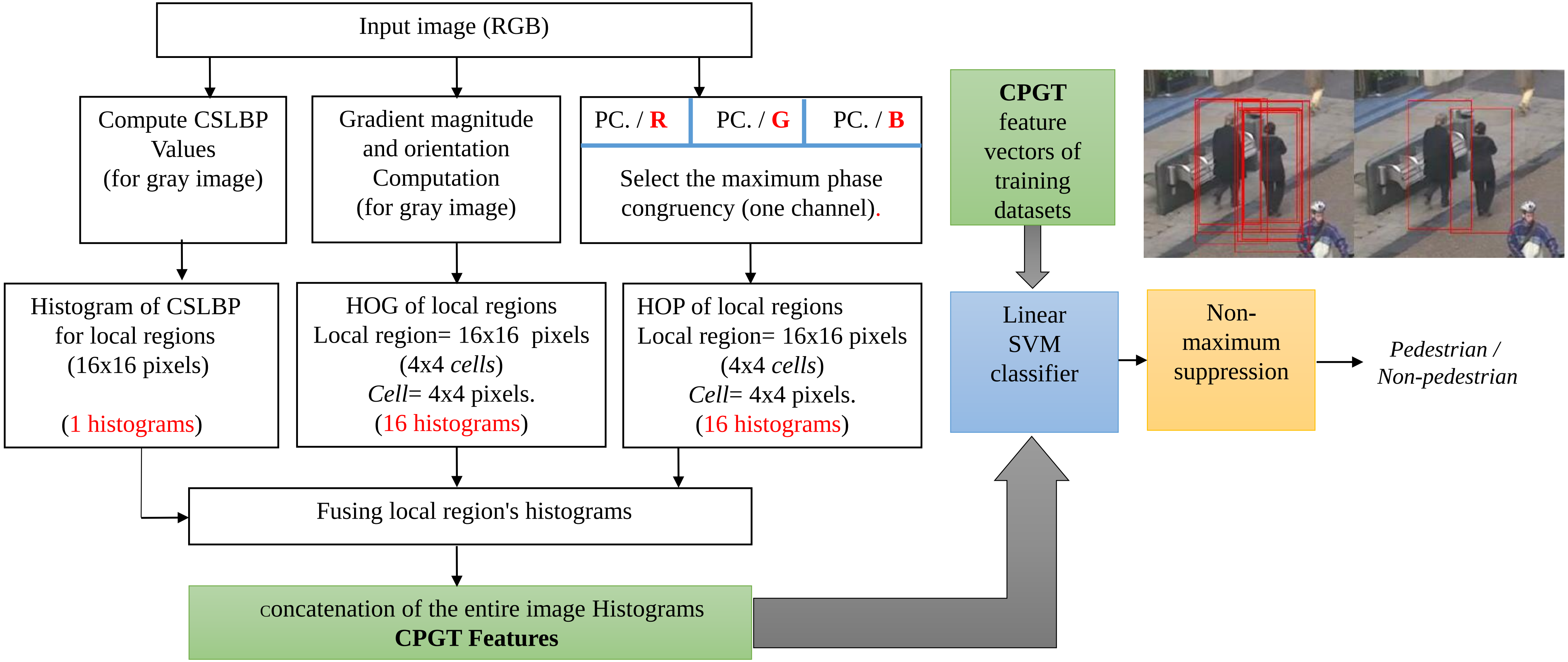


Human Detection



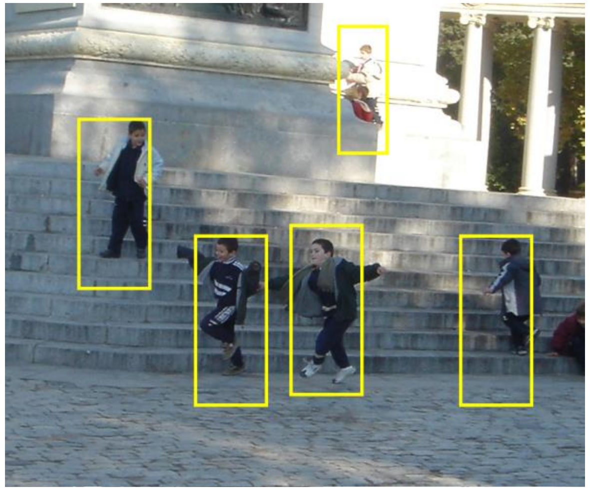
Framework of the Human Detection System

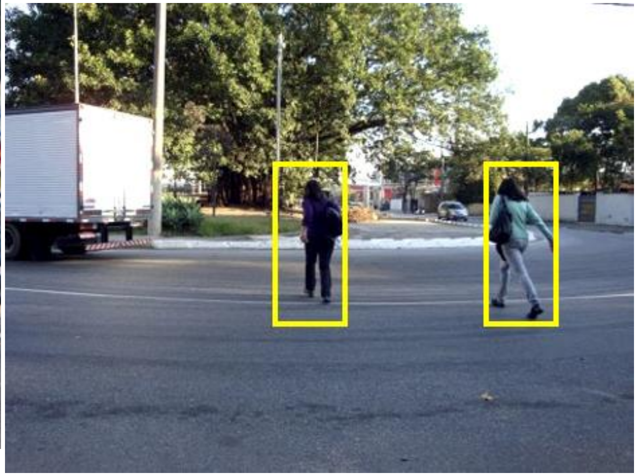
Chromatic domain phase features with gradient and texture (CPGT)

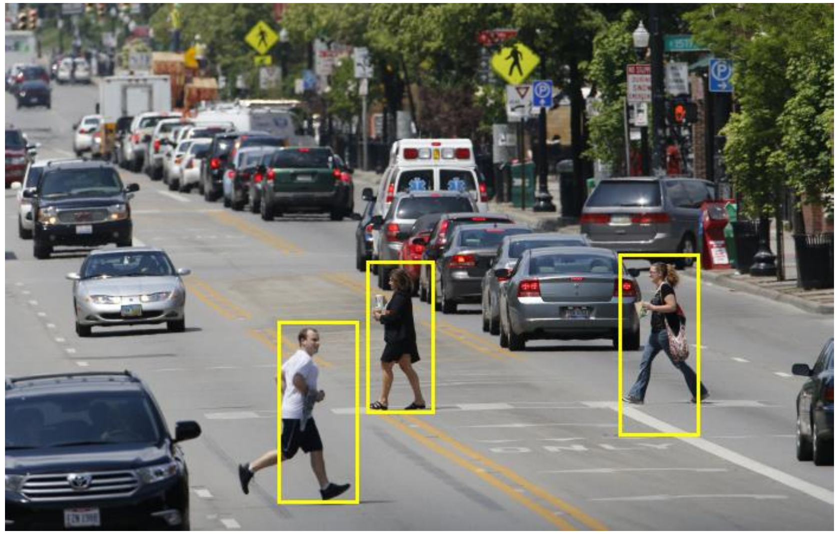


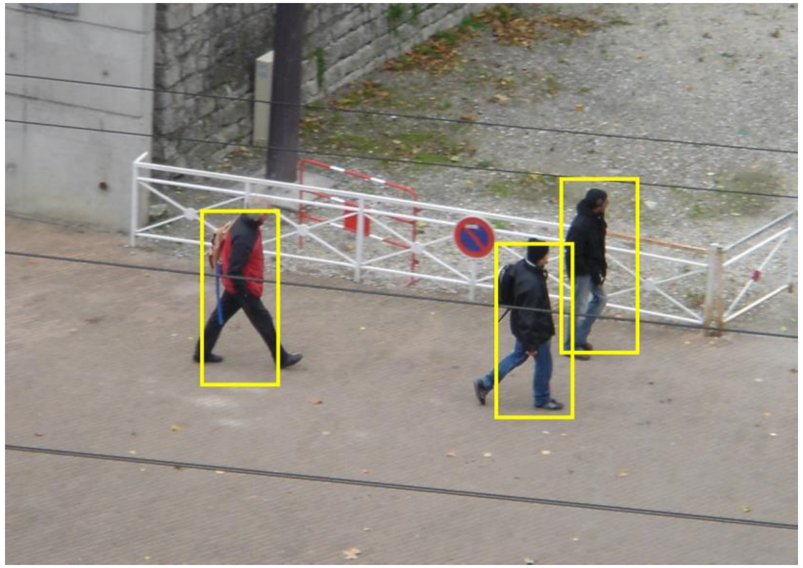
* *HOP= Histogram of Oriented Phase*
* *HOG= Histogram of Oriented Gradient*
* *CSLBP= Central Symmetric Local Binary Pattern.*

CPGT Detector Results

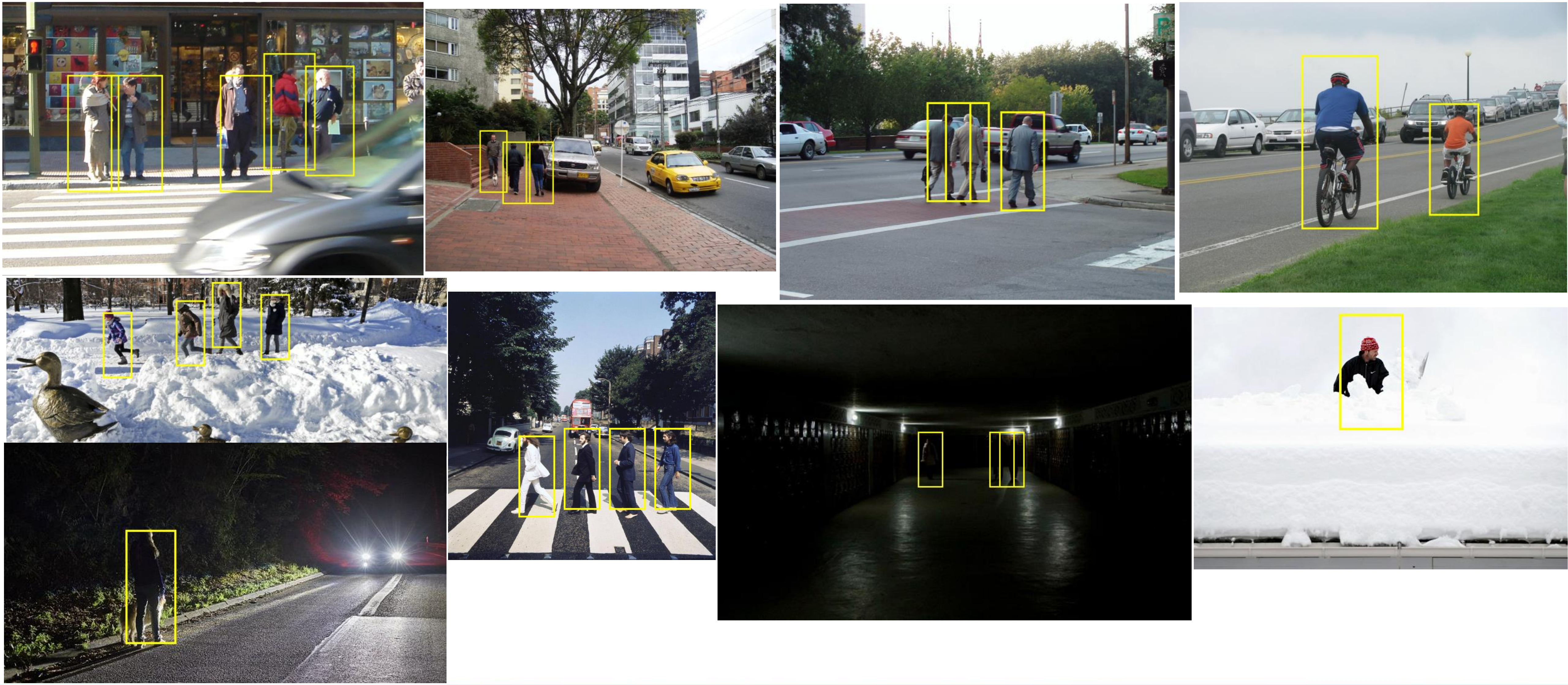








CPGT Detector Results



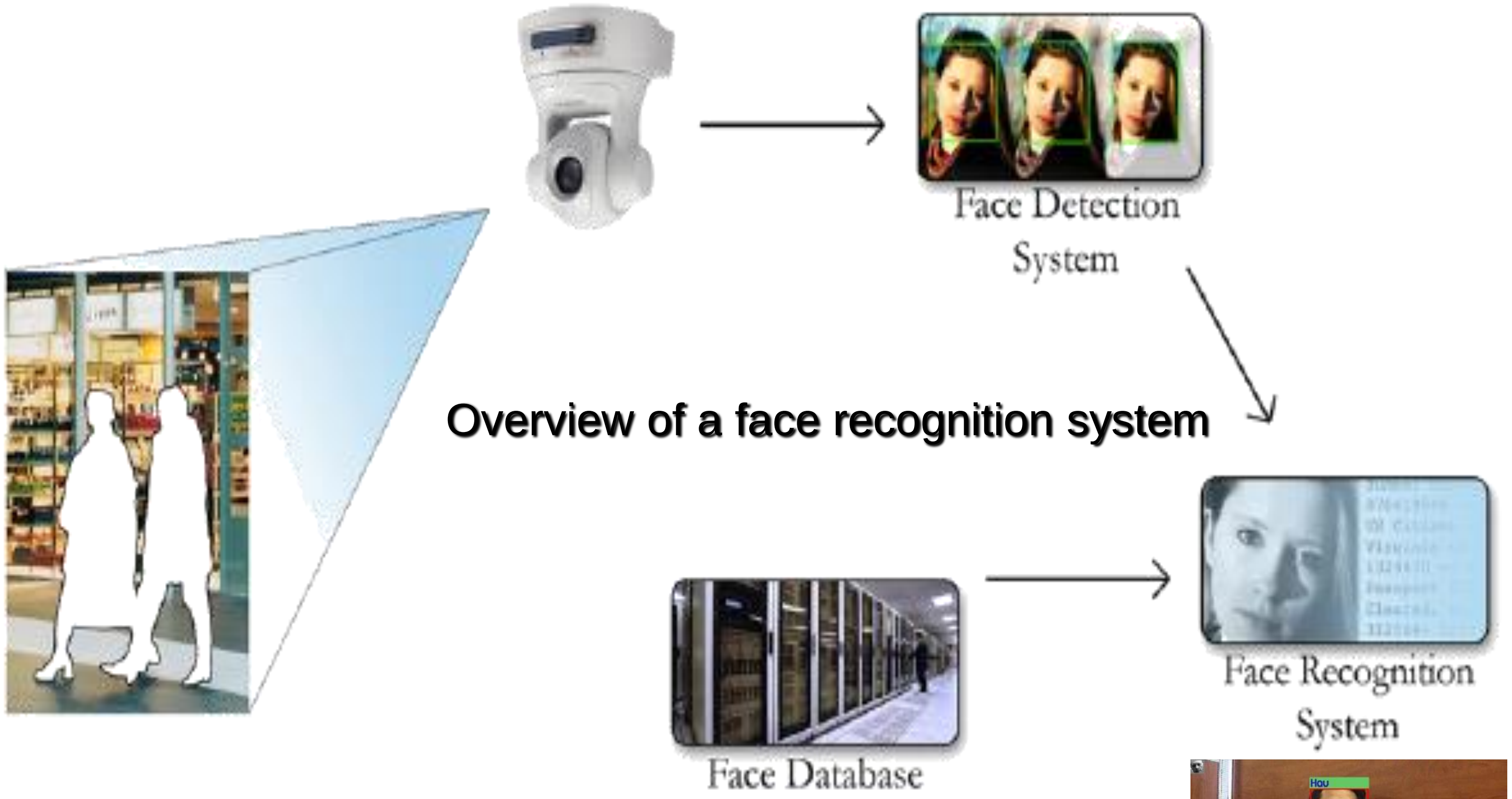
Face Recognition System

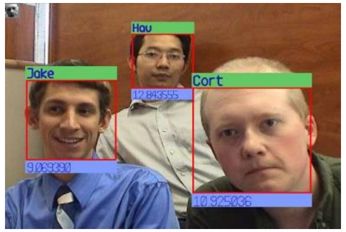
Face Detection – quickly and efficiently locates all faces in a given image region.

Face Features– calculates unique features of each person in the face database that can be used for accurate classification.

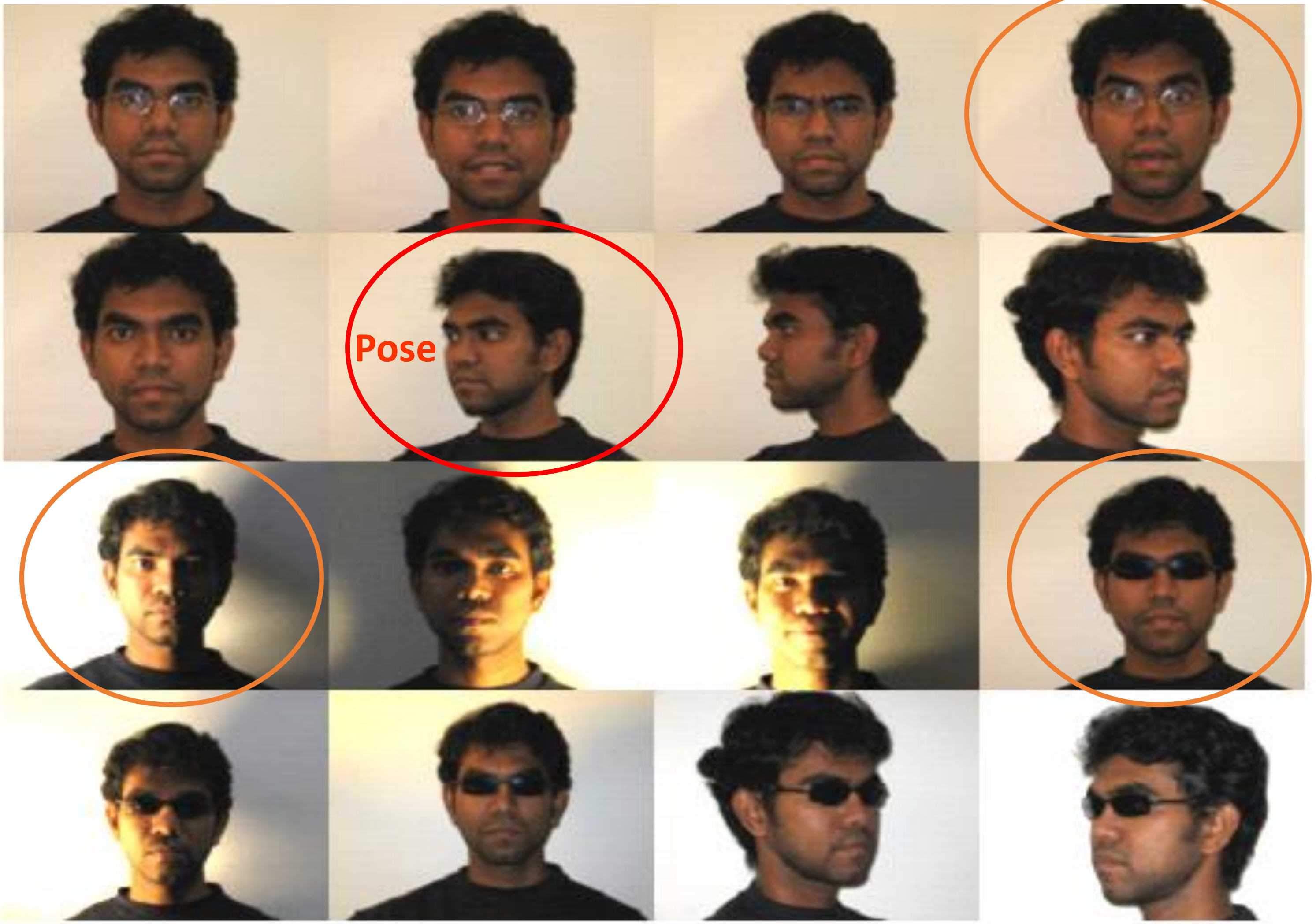
Feature Classification – compares features of face regions obtained from the detection process with face feature data computed from the training stage to determine the identity of individuals.

Face recognition in video



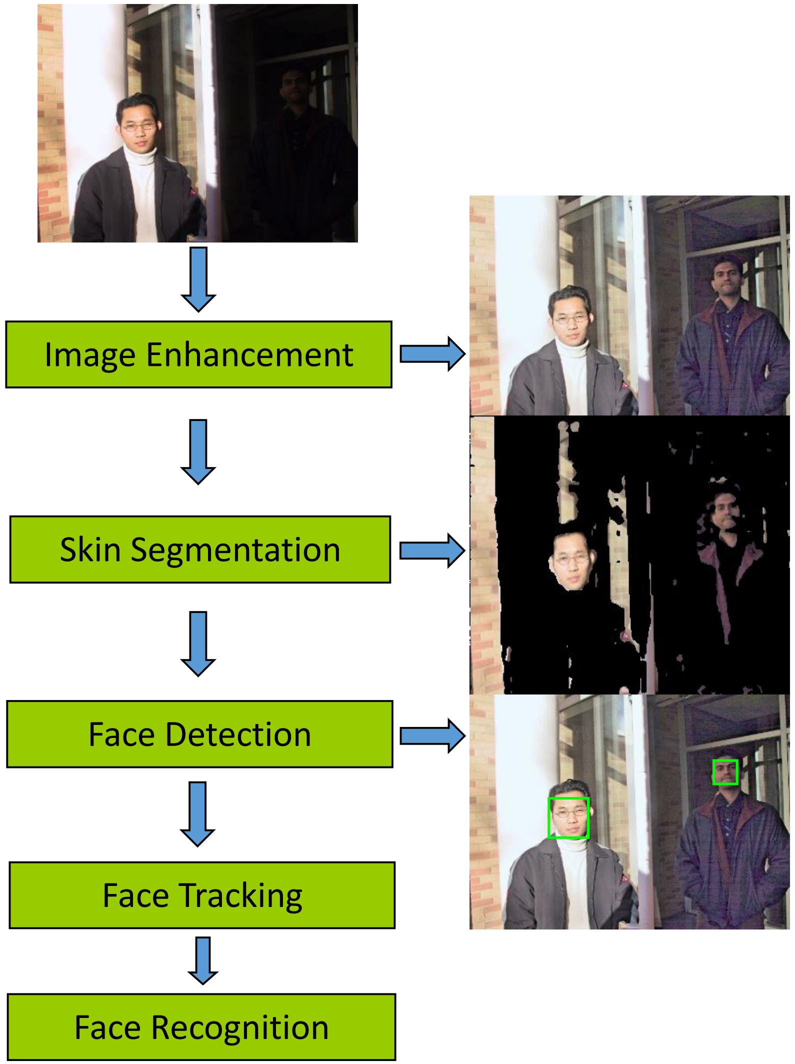


Face Recognition: Appearance Variations



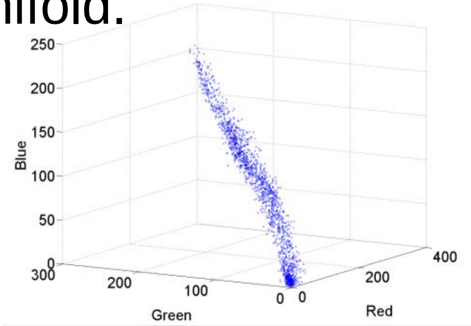
Face Recognition – UD Research

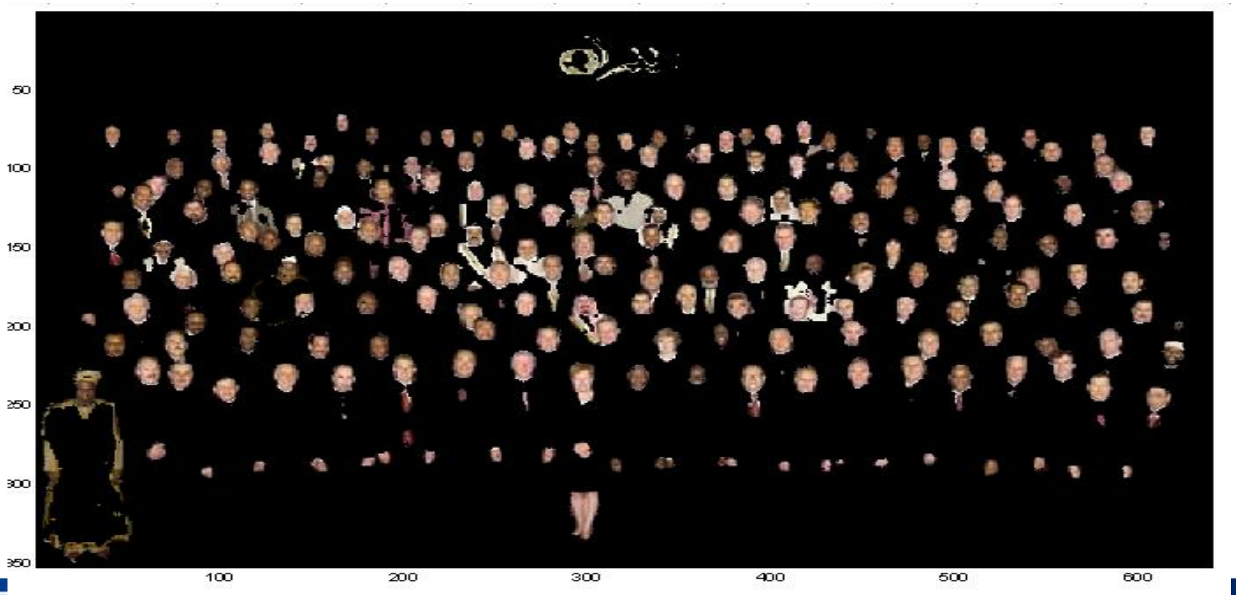
- Images at various lighting conditions are enhanced to a uniform lighting environment.
- In order to reduce the search space for faces in an image frame, the human skin regions are extracted using the color information.
- Search for faces in all skin regions by using a feature matrix developed by a training process.
- Detected faces are tracked in consecutive frames by statistical analysis performed using the concept of particle filter.
- Manifold learning technique for face recognition.



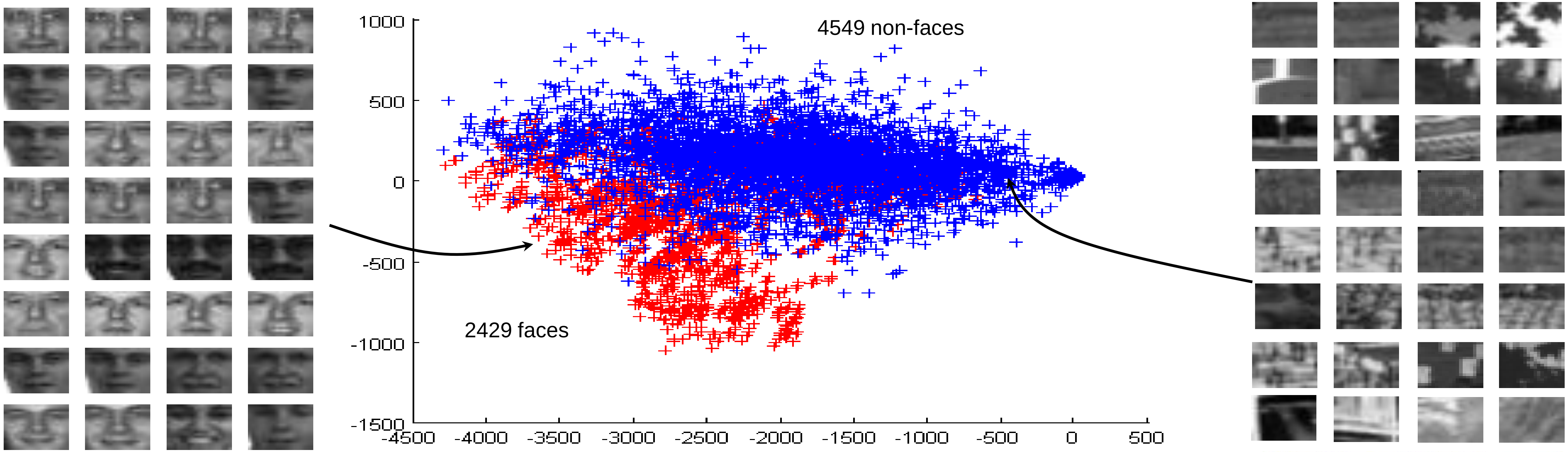
Skin Color Segmentation

Skin colors are forming a nonlinear pipe in the RGB space. It is possible to describe the skin color mathematically using the nonlinear manifold.





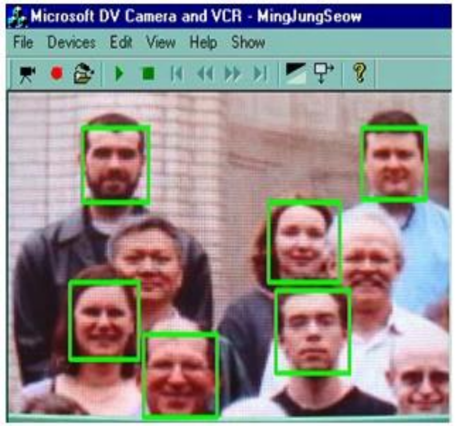
Face Detection



Training with faces and non-faces.

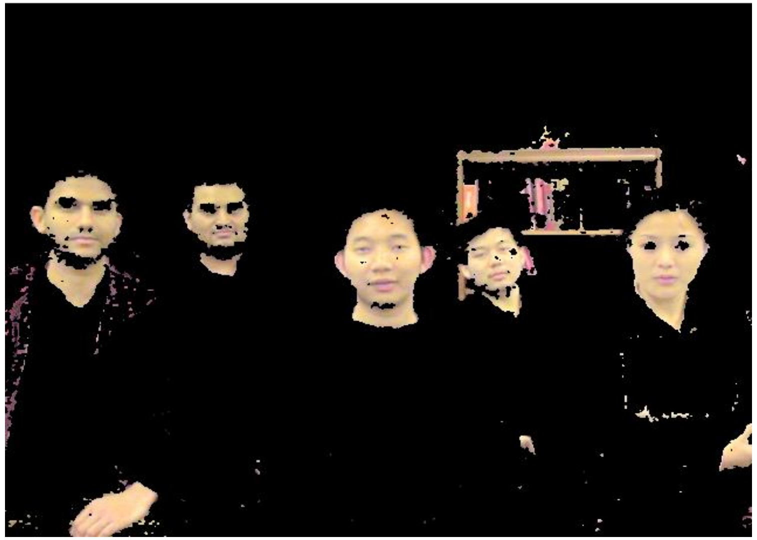
Dimensionality reduction.

Classification.



Face Detection in Enhanced Images

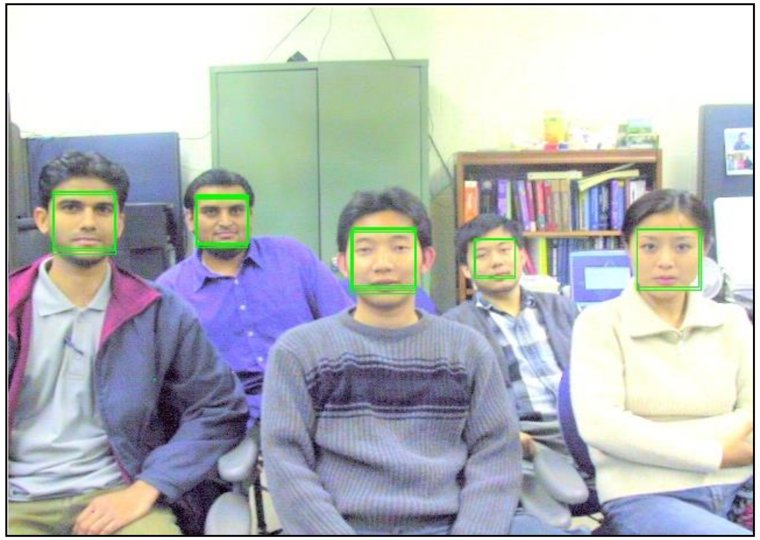
Original image



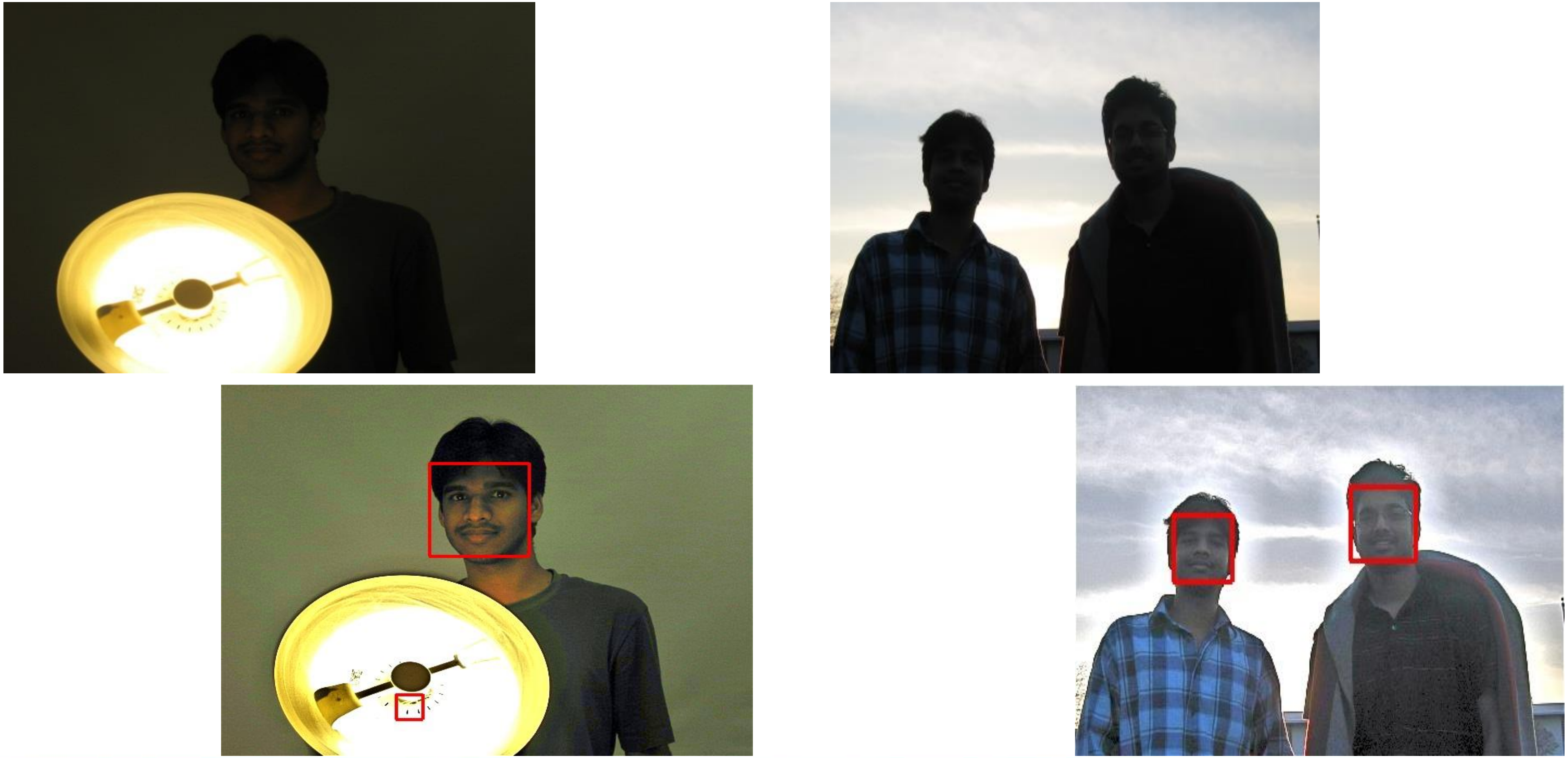
Skin segmented image

Enhanced image

Detected faces



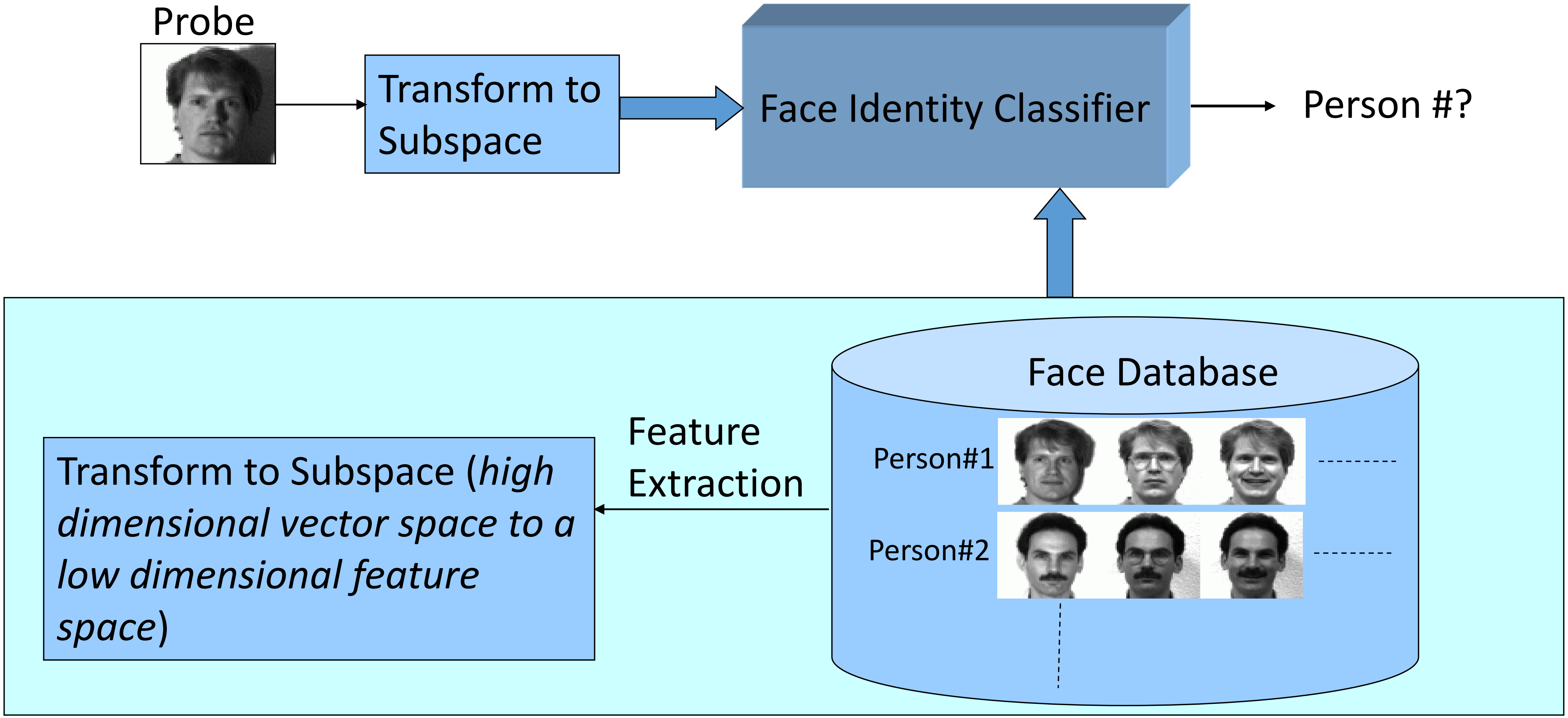
Lighting Invariant Face Detection



Pose Invariant Face Detection

Pose Invariant Face Detection

Face Recognition: A Modular Approach

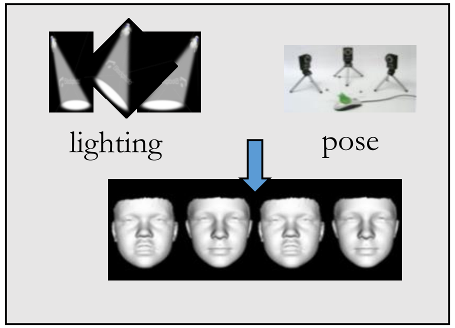


Face Recognition: Object Pose and Orientation Variation

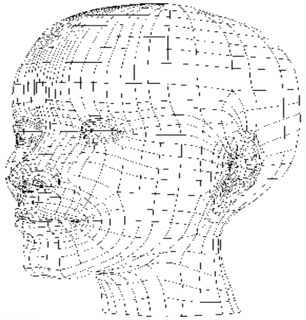
Face images are from UMIST face database

Synthetic Database using Single Training Image

Original image



3D Face Model



45 Degree Side-Lighting from both sides

Top-Lighting Overhead

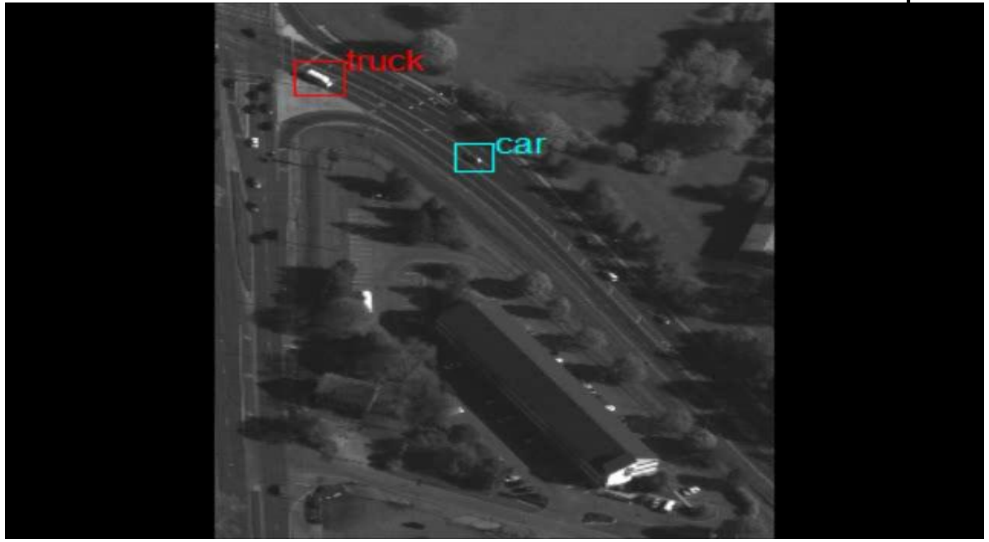
Top-Lighting Overhead with
45 Degree Side-Lighting from both sides

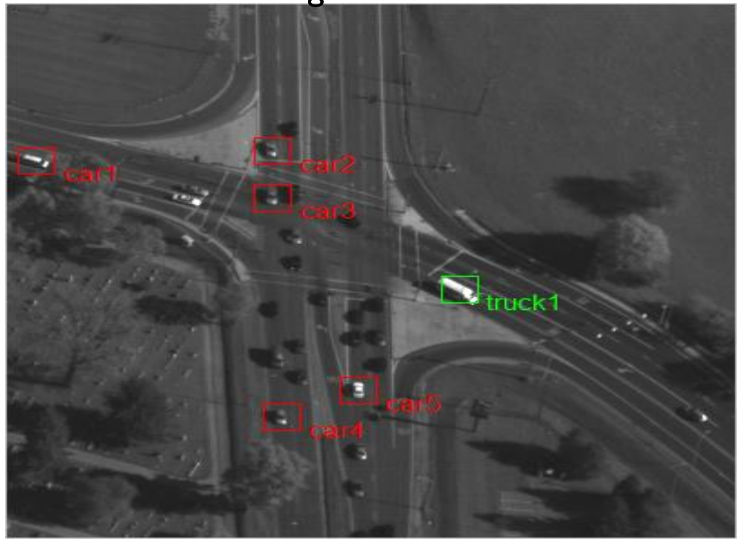
Generated synthetic 2D images

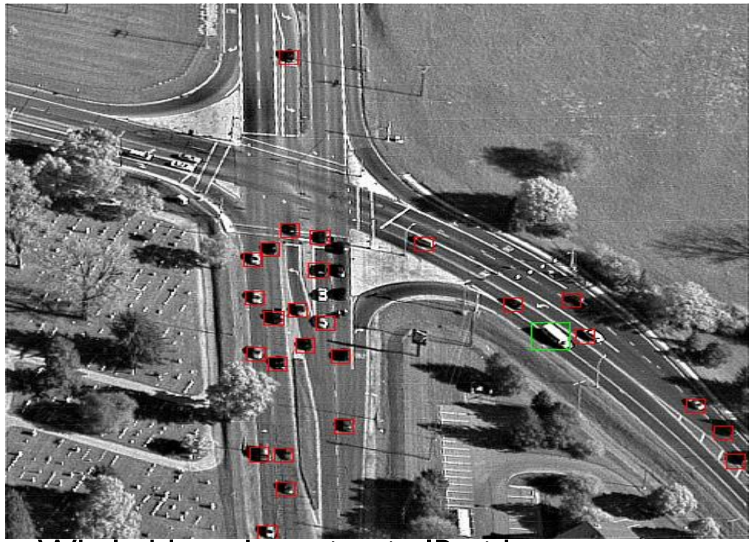
Face Recognition – Moving Forward!

Object Detection, Tracking, and Identification: Wide Area Motion Imagery Data and IR Data Analysis

Object detection and tracking on WAMI data



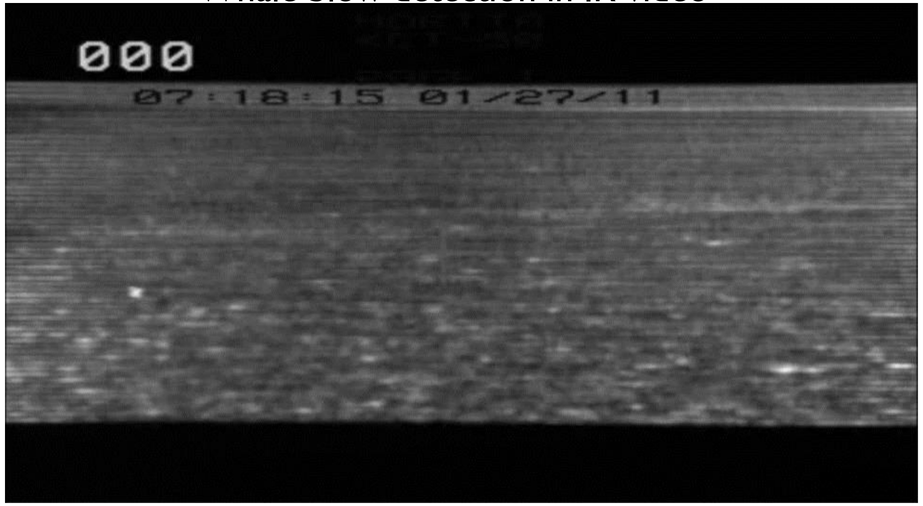




Pedestrian tracking

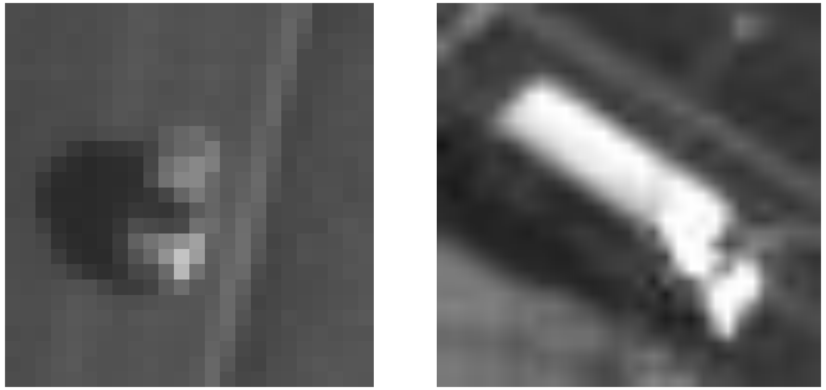
Small boat detection

Whale blow detection in IR video



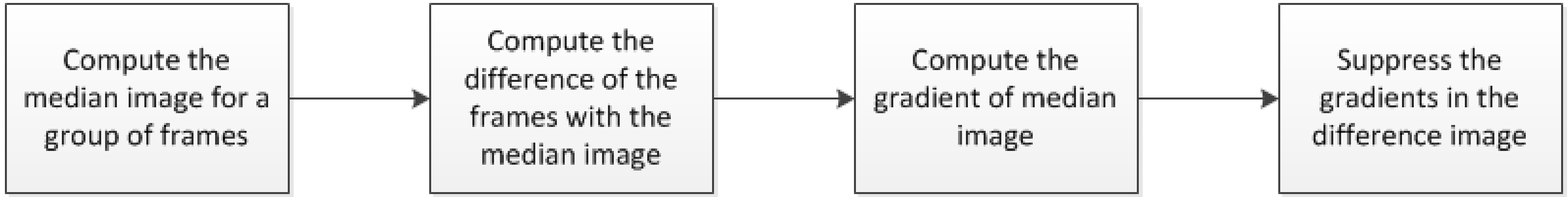
Wide Area Aerial Imagery Data Analysis

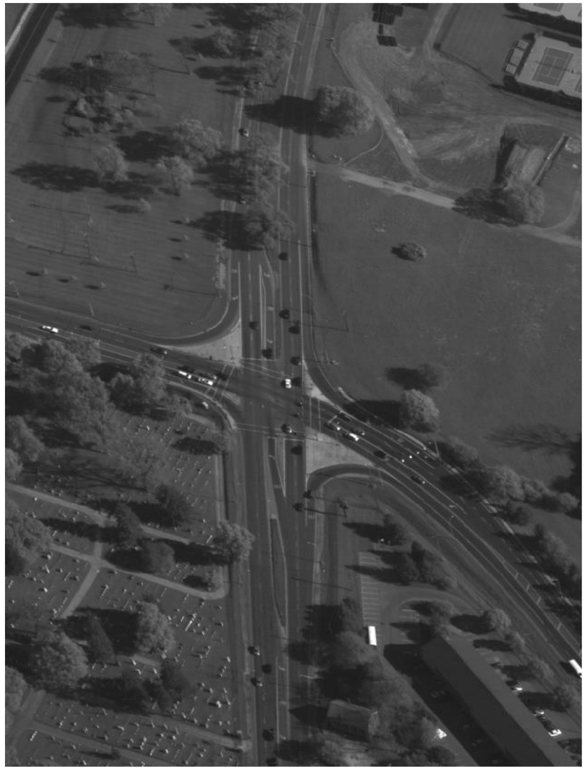
- CLIF – Columbus Large Image Format.
- Data from electro-optic sensors mounted on an aerial platform flying at 7000 feet.
- Six cameras with partially overlapping fields of view.
- Frame size: 4008×2672 pixels at 2 fps.

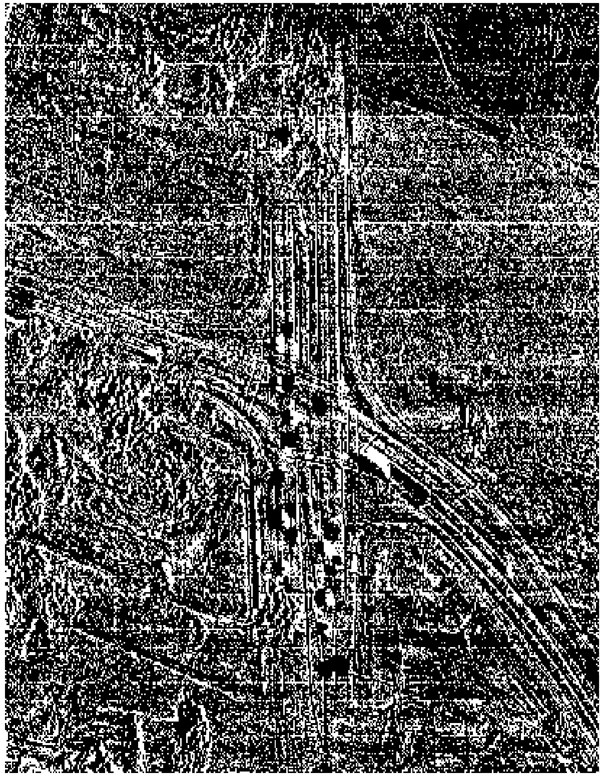


Objects of interest – cars, vans, trucks

Moving Object Detection



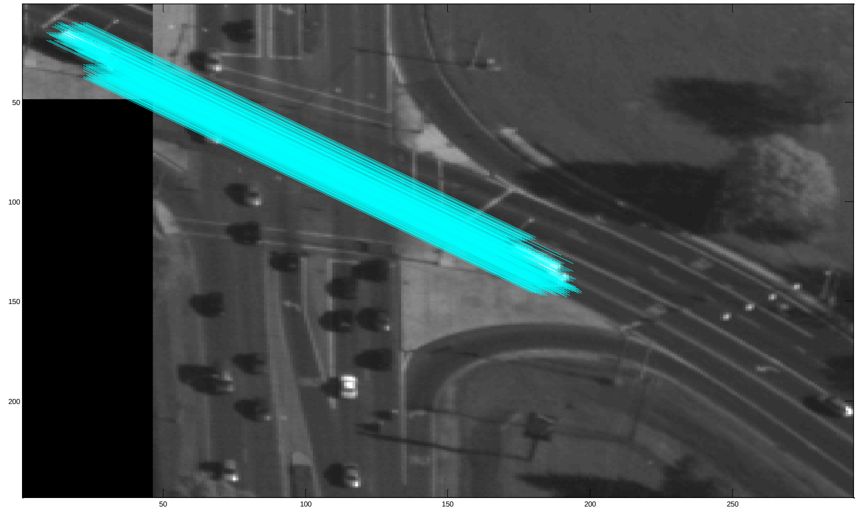

Original frame

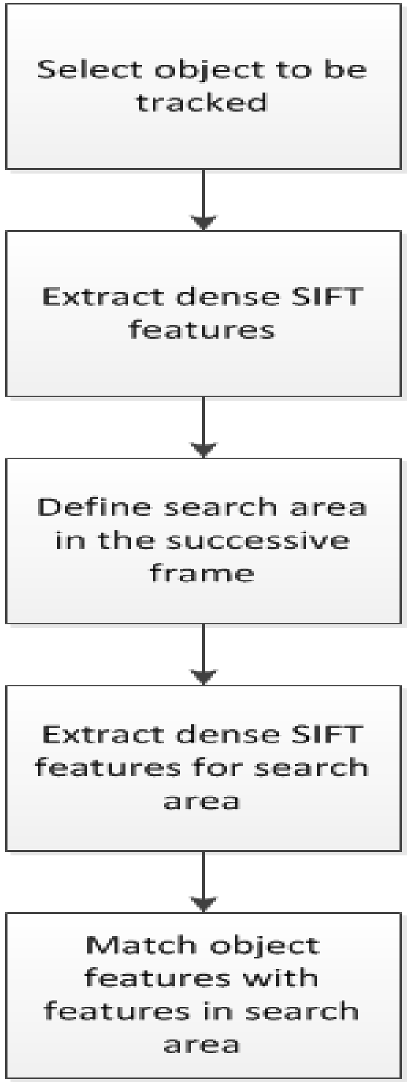

Difference image


After gradient suppression

Object Tracking

- Feature tracking using **Dense SIFT**
  - Extract SIFT features for every pixel .
  - Dense feature set gives a better representation of the object.
  - Matching is based on the criteria that ratio of distances to first and second closest match should be greater than a particular threshold.





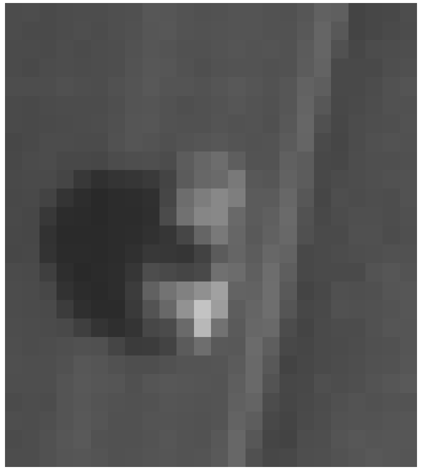
Vehicle Tracking

Tracking multiple objects in a scene with enhancement

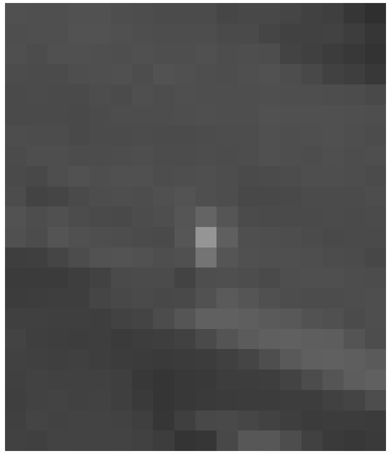
Object Tracking with Enhancement and Super-resolution

Pedestrian Tracking

Track pedestrian movement in long range data (CLIF data)



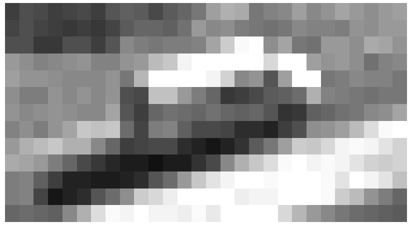
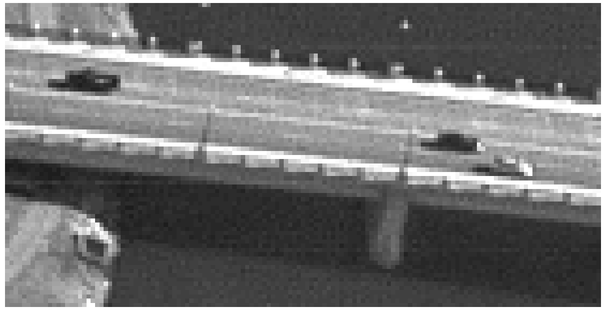
Car
15x15 pixels

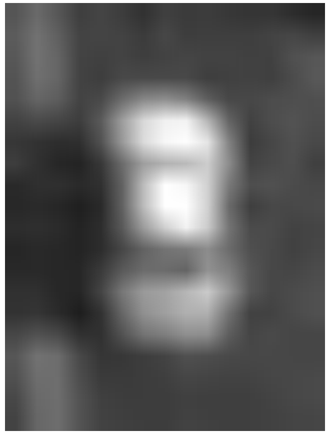


Pedestrian
2x7 pixels

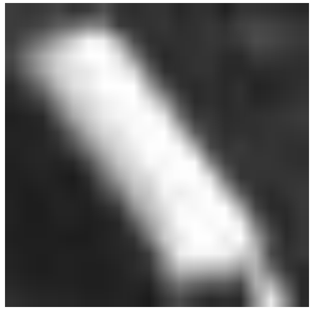
Pedestrian Tracking

Classification Problem on CLIF Data

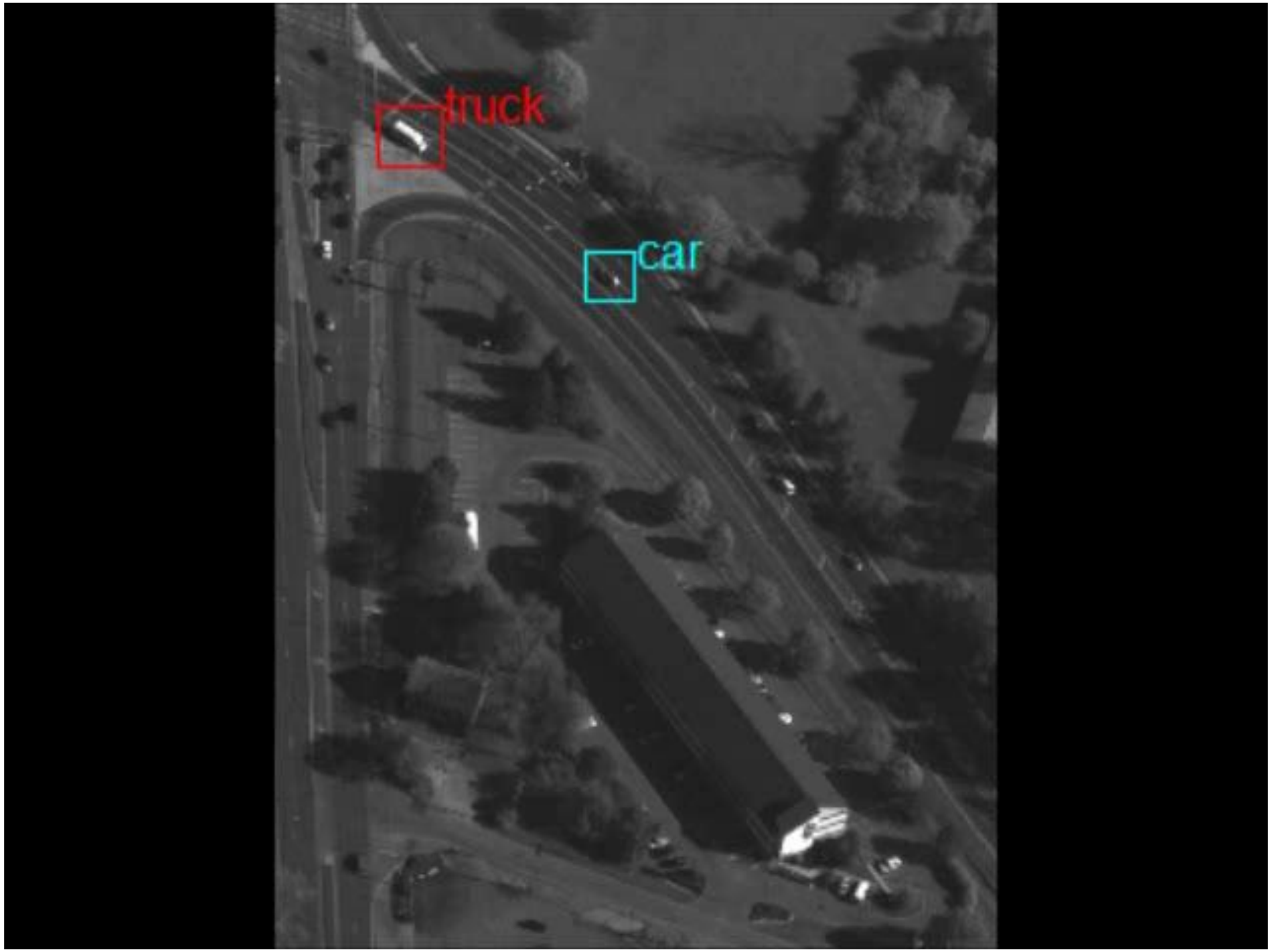


- Low resolution
- Poorly defined contour
- No color information

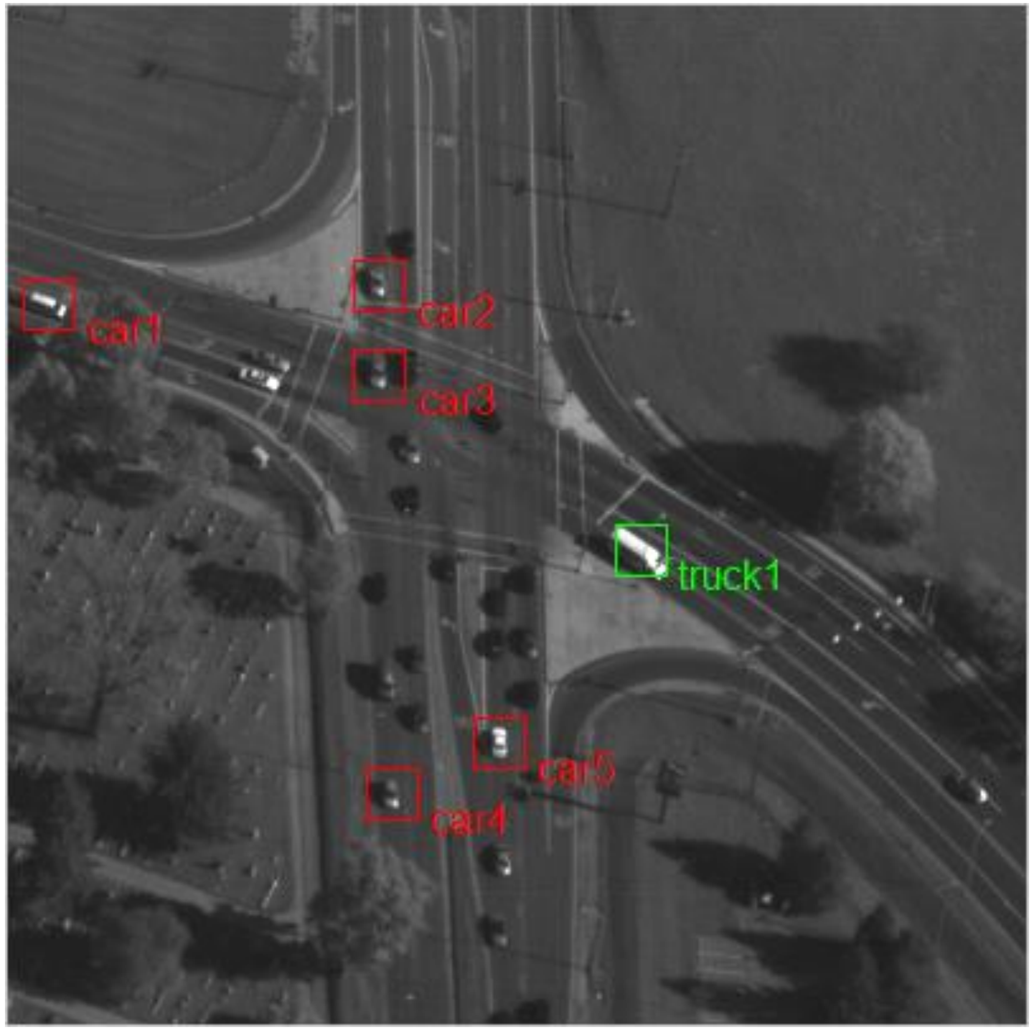


- Trucks and cars: Intensity distributions are significantly different
  - Enhancement is an important preprocessing step
  - Some fuzziness in the intensity distribution
- Classifiable with Linear SVM

Tracking with Classification



Moving Object Classification



Detecting and classifying moving targets into two classes.

Moving Object Classification with Enhancement

The number of detections significantly improves with super-resolution and enhancement.

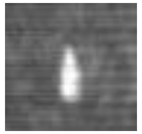
Whale Blow Detection in IR Video: Objective

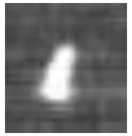
- Detect and track movement of whales during migration
  - Detect presence of whales by detecting whale blows
  - Estimate pod size using timing constraints of whale blows
  - Track whale movement based on their characteristic movement patterns

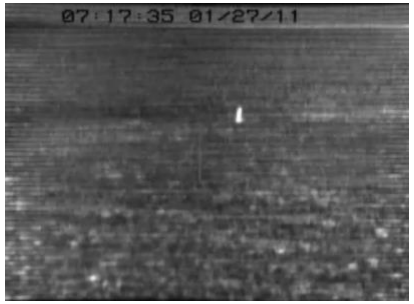
IR Video: Frame Size: 340 × 280 pixels, **Frame Rate:** 30 fps

Characteristics of Whale Blows

- Blow appears as a distinct change in the environment.
  - Whale blow is brighter than the background.
  - Distinctive shape when the blow is full-size.
- Two whale blows will not have same base.
  - Presence of significant distance between two whales.
- Temporal characteristics of the blow.
  - Rise period and fall period.
  - Characteristic variation in blow shape.





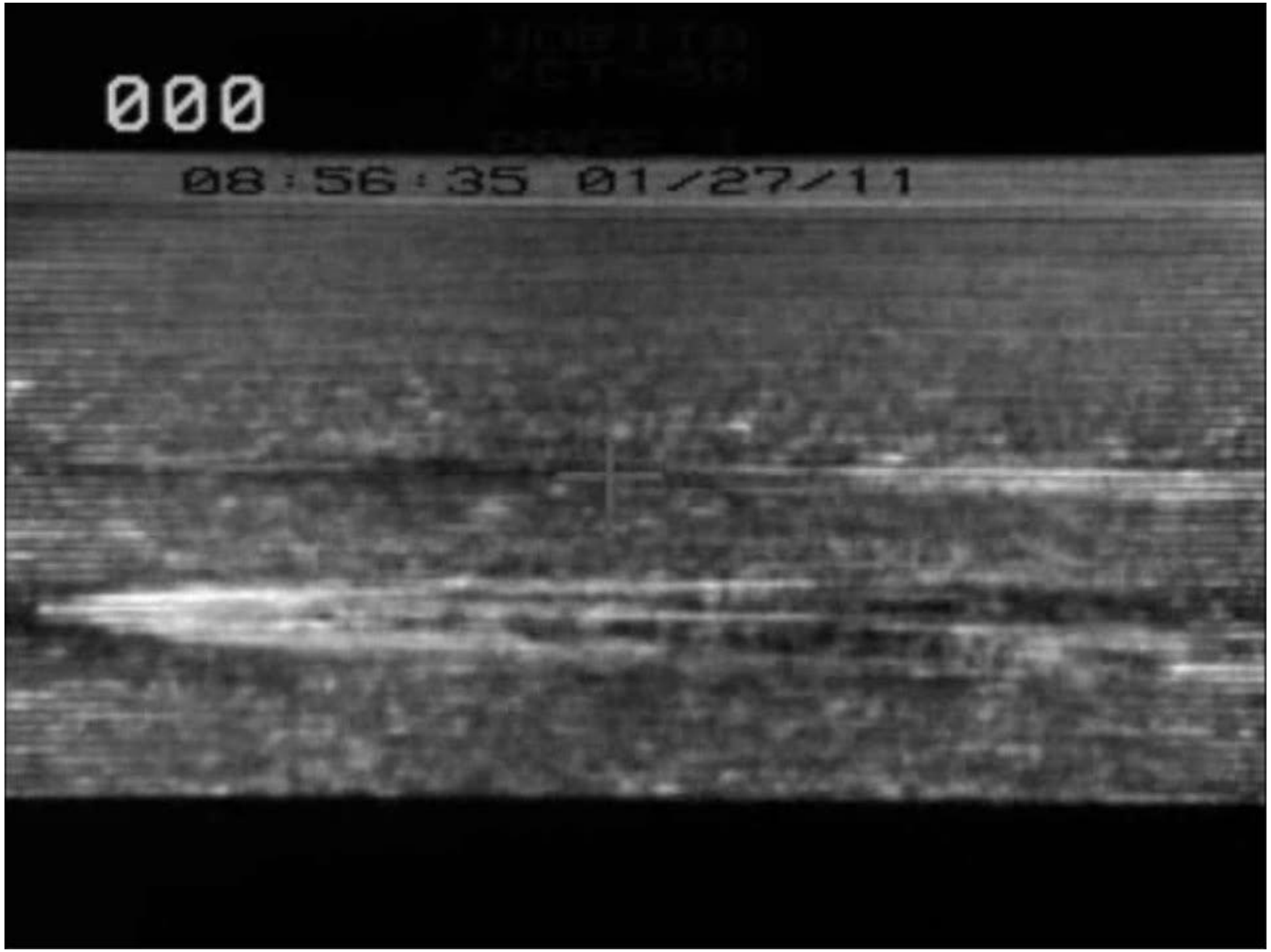


Whale Blow Detection

Video with whale blow

Whale Blow Detection

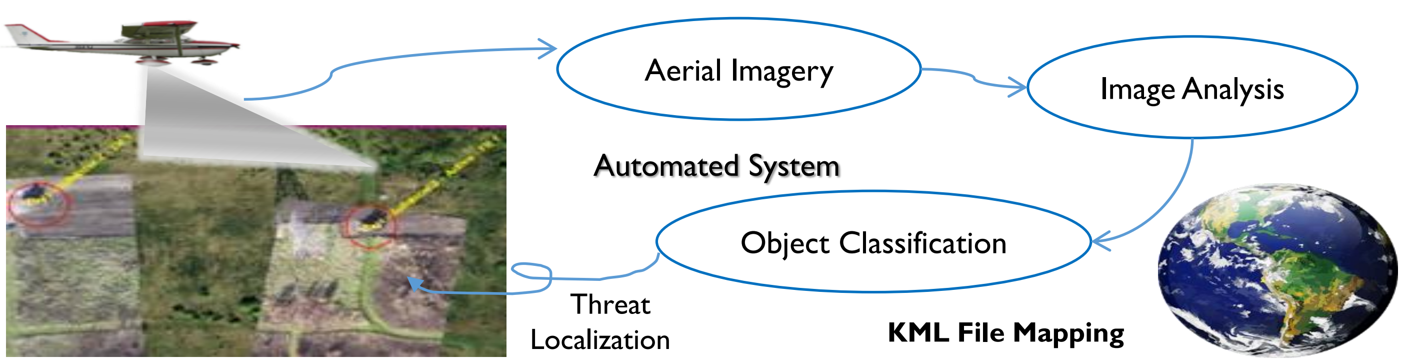
With textural variations on the surface

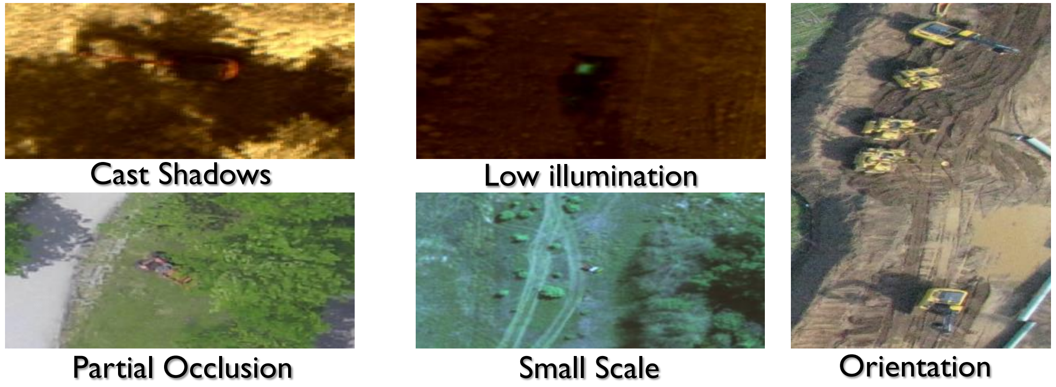


Whale Blow Detection

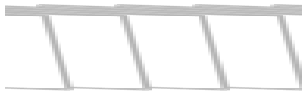
Multiple whale blows

Oil/Gas Pipeline Right-of-Way Automated Monitoring for Pipeline Encroachment and Machinery Threat Detection

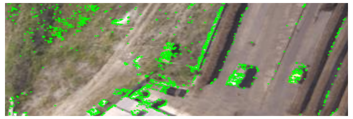




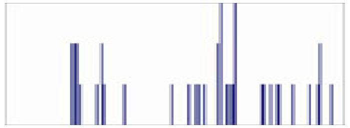
Color Transformation and Enhancement



Key Frame Selection



Key Region Localization

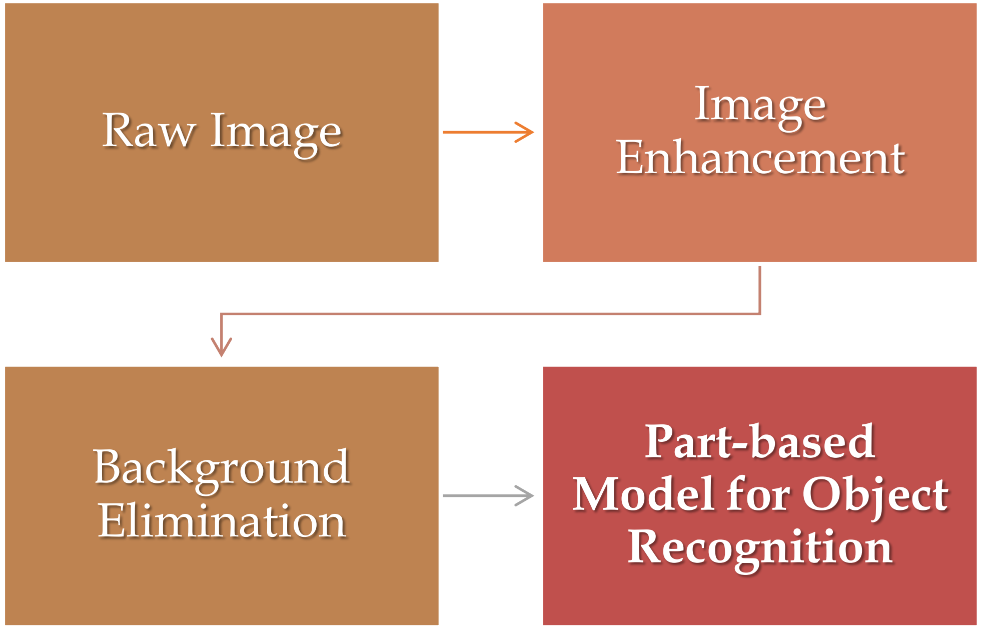


Local Feature Extraction

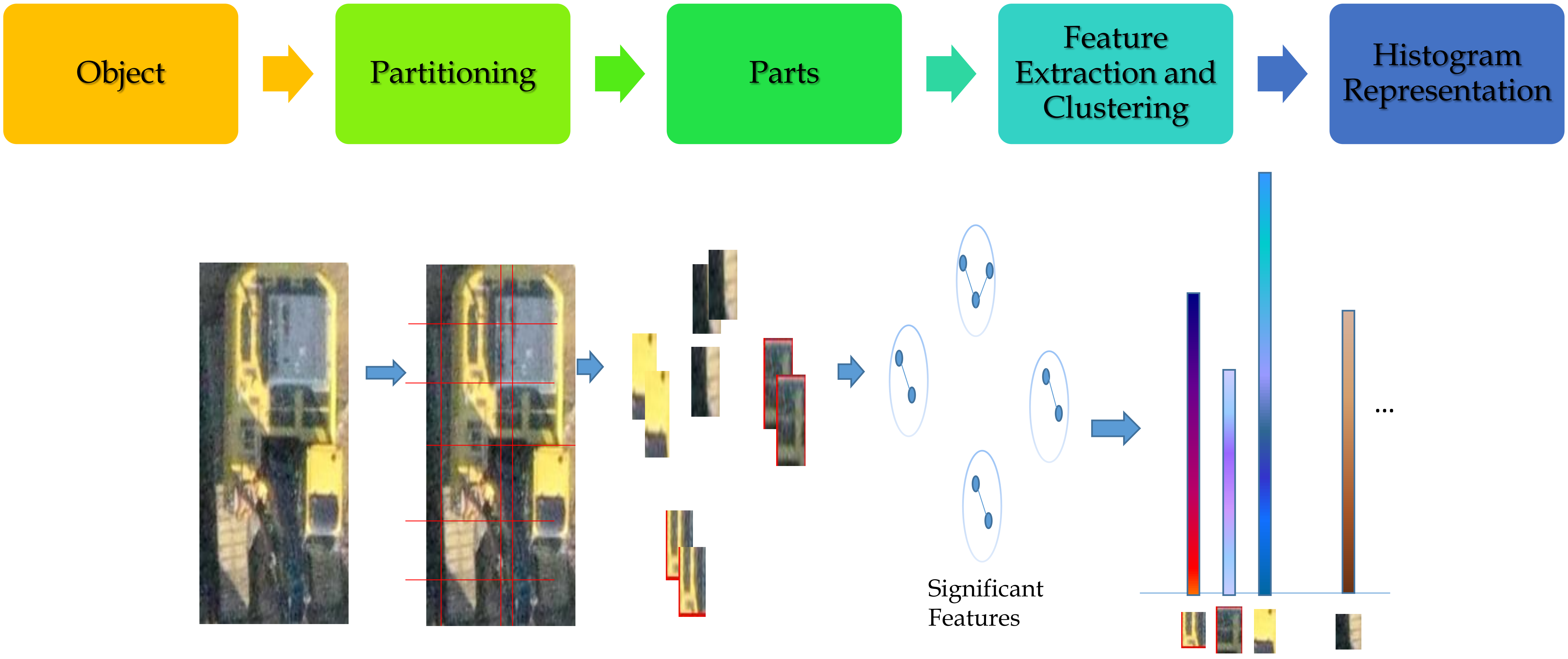
Object Detection and Identification

Part-based Model for Robust Classification

- The purpose of developing a part-based model is to cope with partial occlusion and large appearance variations.



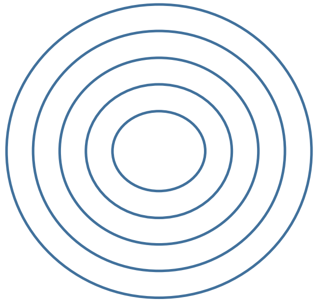
Part-based Model for Robust Classification



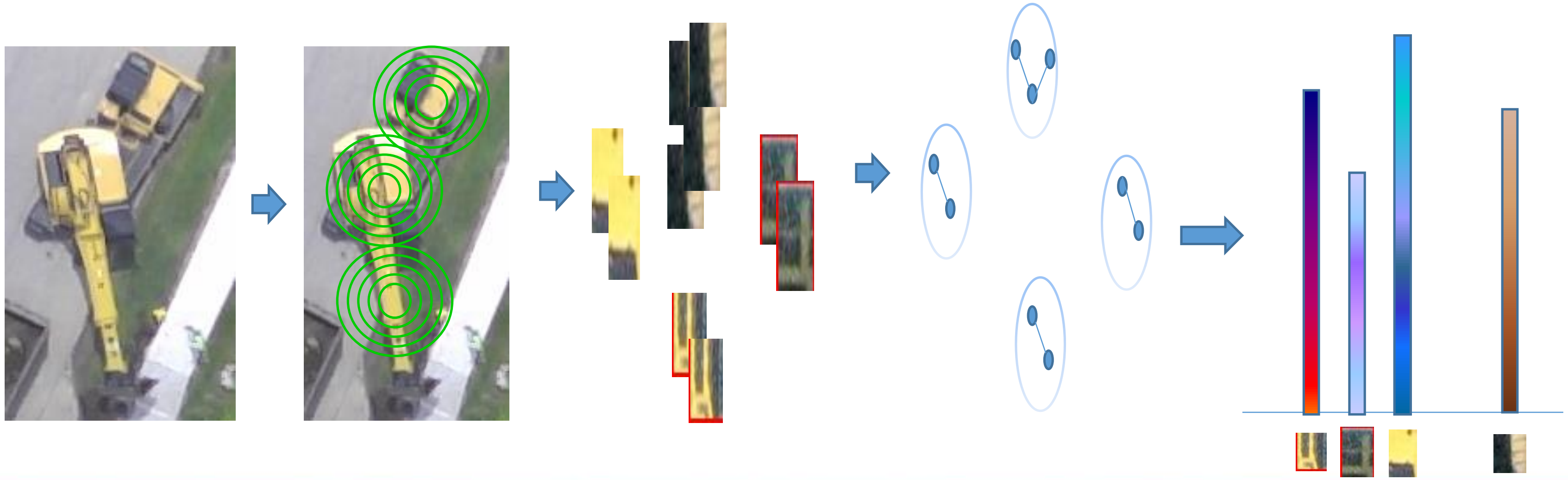
Ringlet Part-Based Model

Method: Using Ring Histogram for each part of objects

- Invariant to rotation
- Still contains spatial information
- Still contains partial occlusion ability



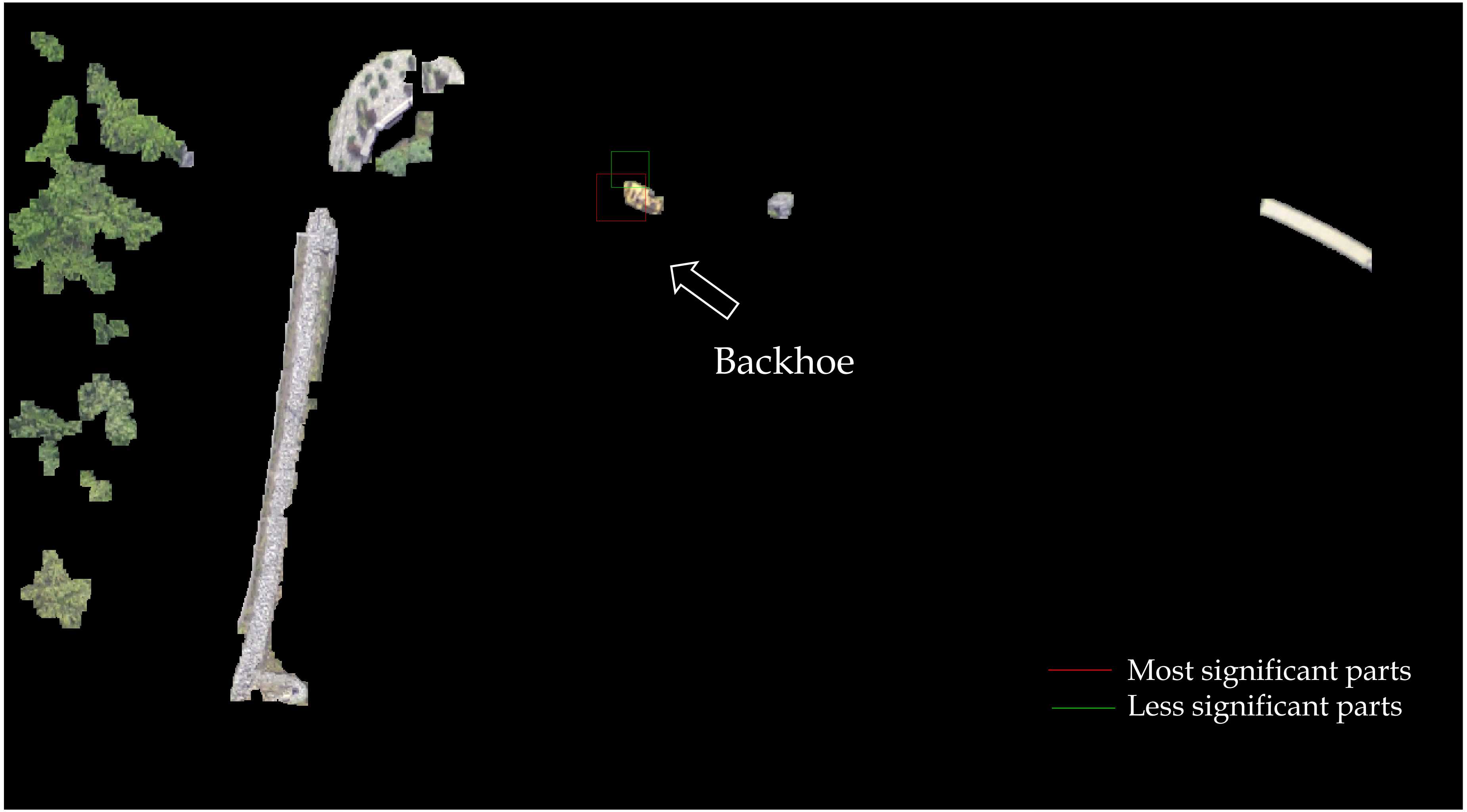
Object → Ring Based Partitioning → Parts → Feature Extraction and Clustering → Histogram Representation



Raw Image– Non Occlusion

Part-based Detection – Non Occlusion

Final Detection Output

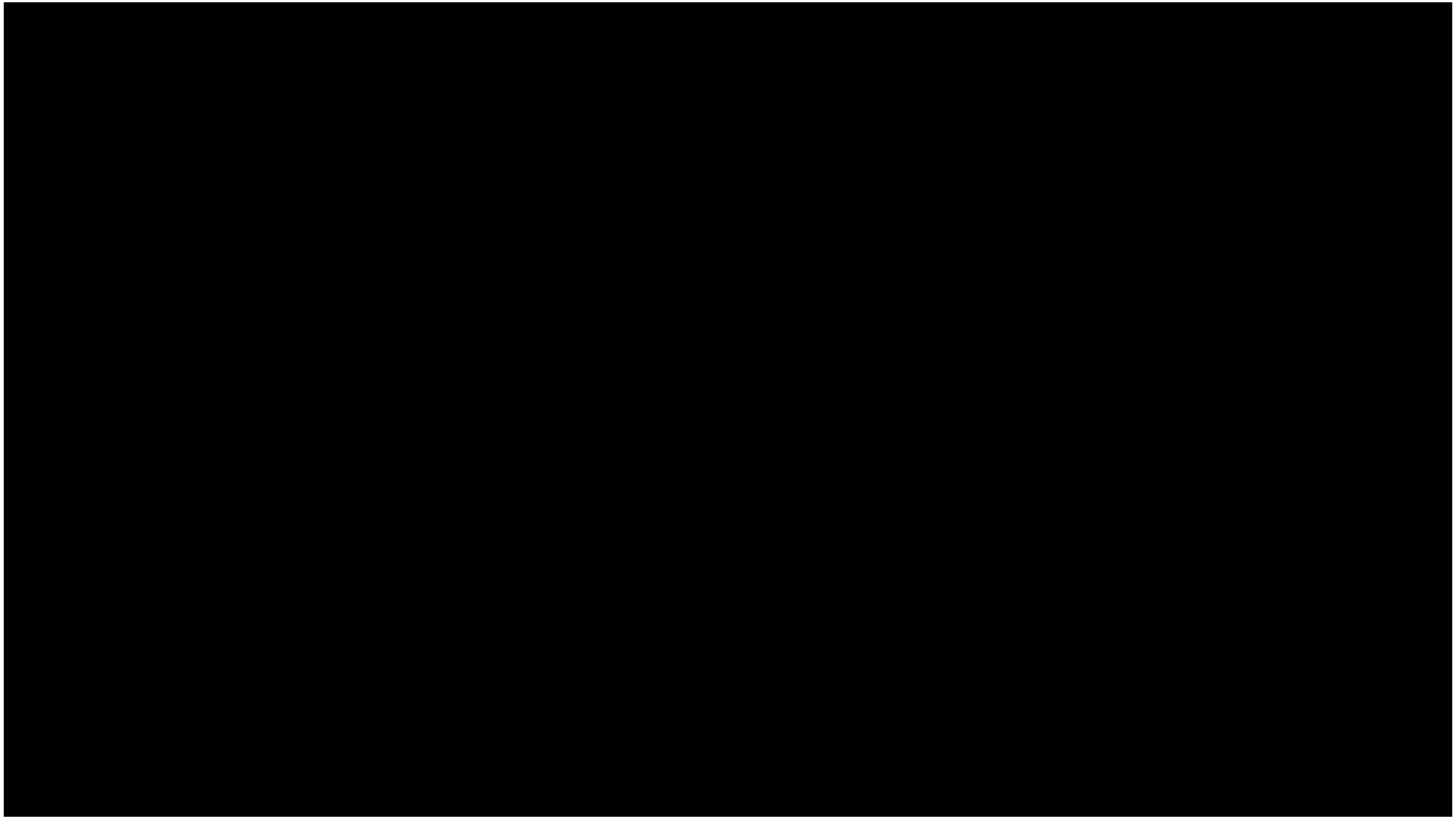


Raw Image– Partial Occlusion

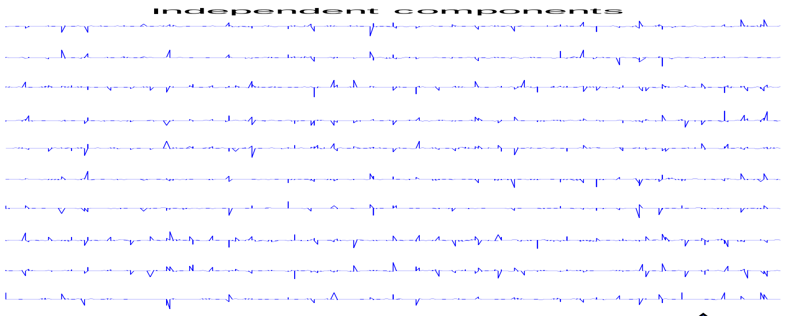
Part-based Detection – Partial Occlusion

Final Detection Output

Threat Detection Results



Brain Signal Analysis:
Emotion Recognition and Brain Machine Interface



Medical Applications

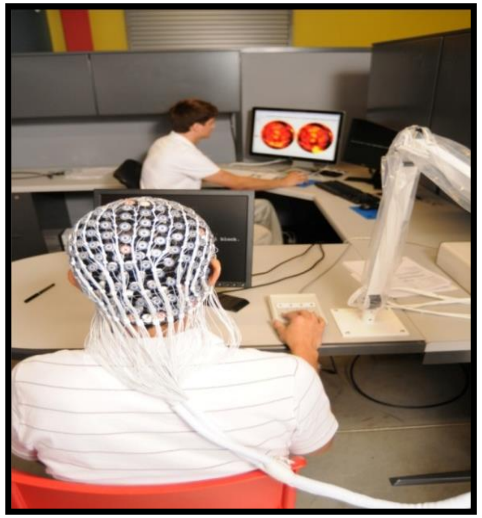
Brain machine interface

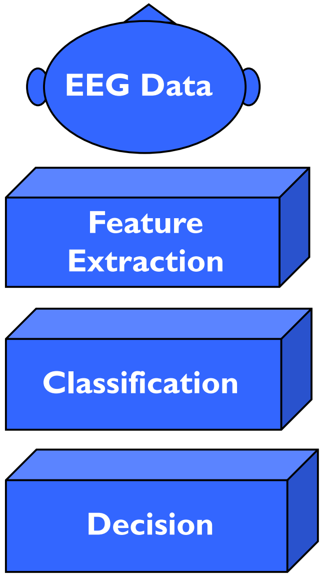
Security

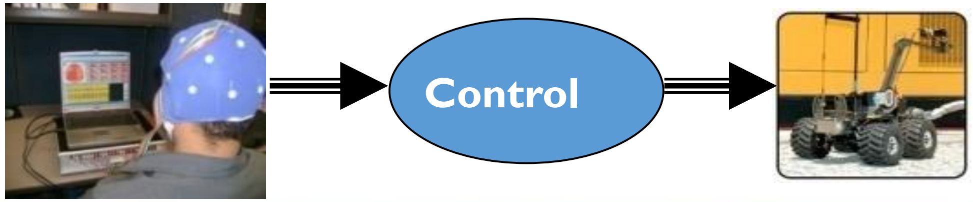
Intentions, motives

Efficiency

Stress detection, fatigue assessment







Thanks

Sensing, Processing and Automatic Decision Making in Real Time