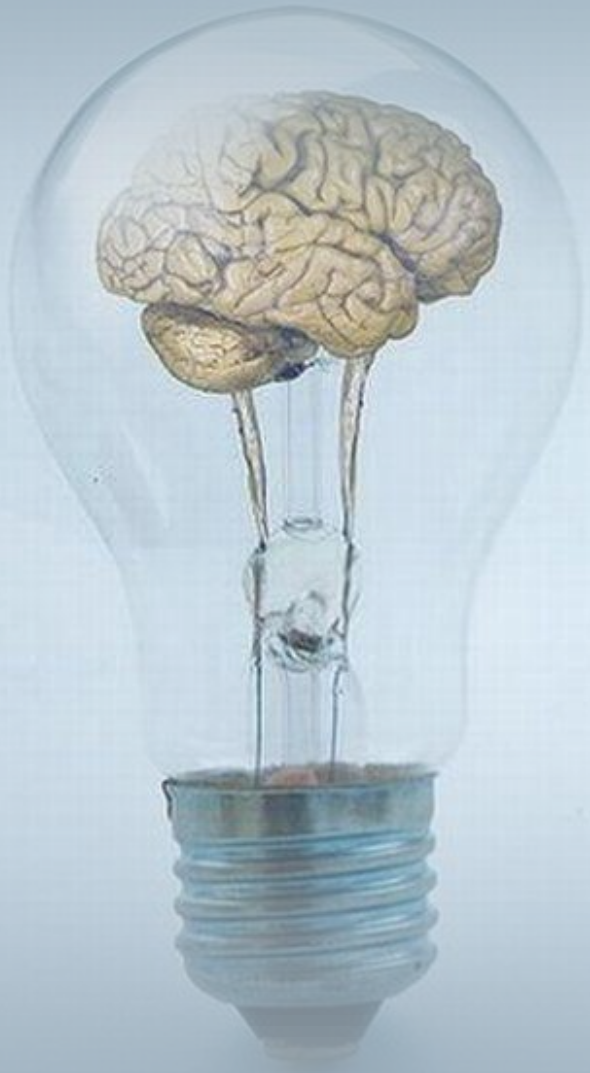


Learning on Manifolds

Dr. Fatih Porikli

August 18, 2010

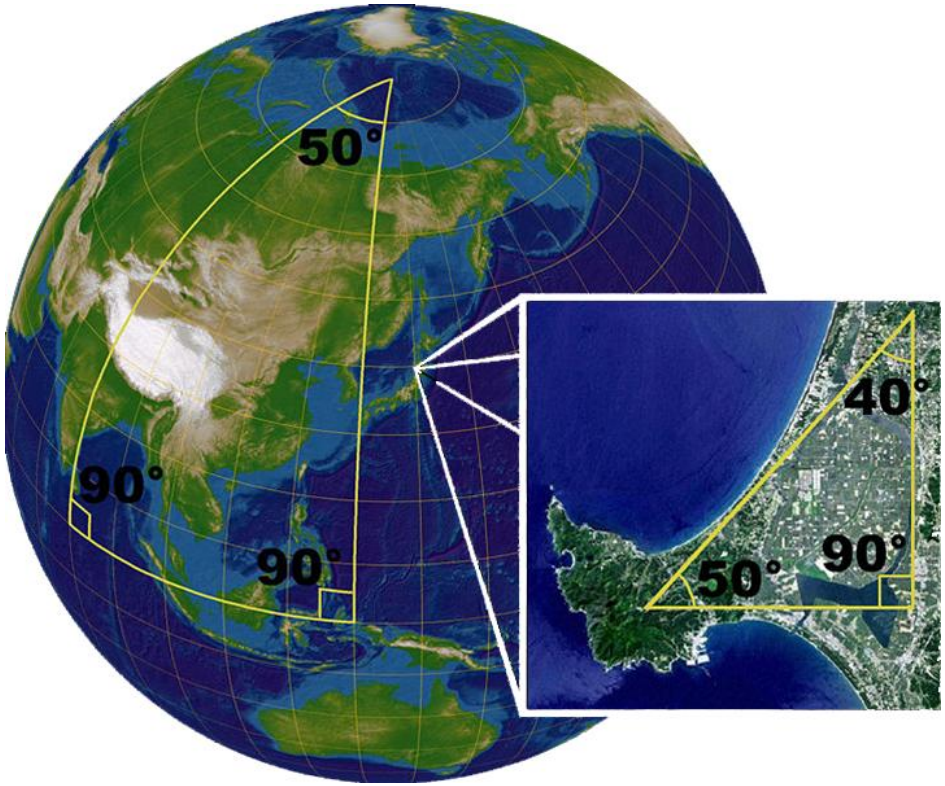
Thanks to Dr. Oncel Tuzel, Dr. Pan Pan, Dr. Peter Meer, Dr. Raghav Subbarao

Manifolds

Manifolds

Manifolds

- A manifold is a “topological space” in which every point has a neighborhood which resembles Euclidean space, but in which the global structure may be more complicated



Manifolds

Manifolds

A group $\mathcal{G}$ is a set that is endowed with a binary operation and satisfies the closure, associativity, identity, and invertibility properties.

Set of integers $\mathbb{Z}$ under addition where the identity is 0

Set of integers $\mathbb{Z}$ under multiplication where the identity is 1

A subset of $\mathcal{G}$ is called as a subgroup if it satisfies all the group properties.

Set of positive rational numbers $\mathbb{Q}^+$ under multiplication

Set of negative rational numbers $\mathbb{Q}^-$

Manifolds

A non-commutative group is the set of invertible $n \times n$ square matrices under matrix multiplication: the general linear group $GL(n)$.

The special linear group $SL(n)$, which is the set of $n \times n$ matrices with unit determinant, is a subgroup.

Manifolds

A topological space is a set $\mathcal{S}$ together with a family of subsets $\mathcal{T}$ if the empty set $\emptyset \in \mathcal{T}$ and $S \in \mathcal{T}$, the union and intersection of any family of sets in $\mathcal{T}$ lies in $\mathcal{T}$.

Sets in $\mathcal{T}$ are called open sets of the topological space.

Any open set $\mathcal{U} \in \mathcal{T}$ which contains point $X \in \mathcal{S}$ is called the neighborhood of the point.

A Hausdorff (separated) space is a topological space in which distinct points have disjoint neighborhoods.

Real numbers constitute a Hausdorff space.

Manifolds

For functions defined on Hausdorff spaces it is possible to introduce continuity:

> As we move towards a point X, the value of the function gets closer to the value of the function at the point.
>
> Being 'close' to a particular point is determined by its neighborhood and the continuity of a function that maps open sets of the topology.

A mapping between two topological spaces is called continuous if the inverse image of any open set with respect to the mapping is again an open set.

Manifolds

A bijective (one-to-one, onto) mapping continuous in both directions is a homeomorphism.

Such mappings preserve the topological properties of a given space.

Two spaces with a homeomorphism between them are called homeomorphic, and from a topological viewpoint, they are the same, e.g. a square and a circle are homeomorphic to each other.

apple orange

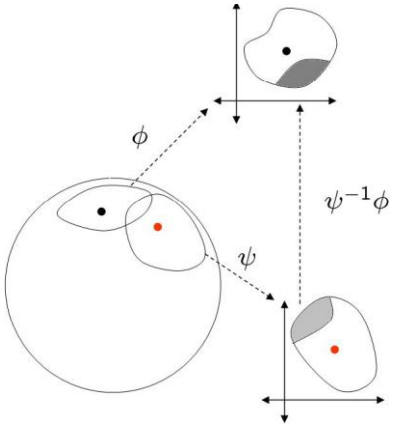
Manifolds

A manifold $\mathcal{M}$ of dimension d is a Hausdorff space for which every point has a neighborhood that is homeomorphic to an open subset $\mathcal{U}$ of $\mathbb{R}^d$.

In other words, a manifold corresponds to a topological space which is locally similar to an Euclidean space.

Manifolds

For any point $X \in \mathcal{M}$, there exists an open neighborhood $\mathcal{U} \subset \mathcal{M}$ containing the point and homeomorphism ϕ mapping the neighborhood to an open set $\mathcal{V} \subset \mathbb{R}^d$, such that $\phi : \mathcal{U} \mapsto \mathcal{V}$. The pair $(\mathcal{U}, \phi)$ is called as a coordinate chart.



An atlas is a family of charts.

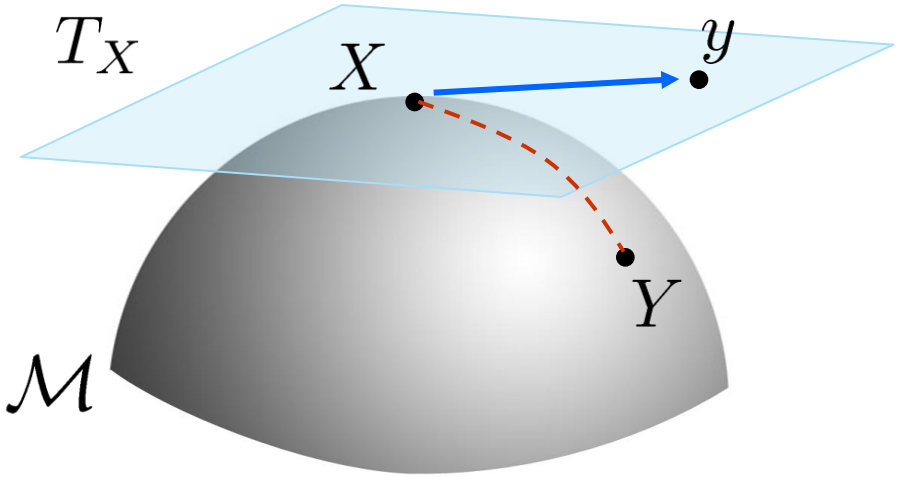
A differentiable manifold C^k is a topological manifold equipped with an equivalence class of atlas whose transition maps are k-times continuously differentiable.

If all its partial derivatives exist, then it is a smooth manifold C^∞.

Manifolds

It is possible to define the derivatives of the curves on a differentiable manifold and attach to every point X a tangent space T_X.

A real vector space that intuitively contains the possible directions in which one can tangentially pass through X.



Manifolds

The tangent space can be thought of as the set of allowed velocities for a point constrained to move on the manifold. Mathematically, it is a generalization of the idea of a directional derivative in Euclidean space.

The tangent space is a vector space, thereby it is closed under addition and scalar multiplication.

Riemannian Manifolds

Bernhard Riemann

1826-1866

Riemannian Manifolds

Riemannian geometry adds further structure to an analytic manifold by defining an symmetric, positive definite bilinear form on the tangent space at each point on the manifold.

A bilinear form on a vector space is a mapping b from $\mathbb{R}^d$ to $\mathbb{R}$ that is linear.

A bilinear form is positive definite if $b(x, y) \geq 0$ with equality occurring if and only if $x = 0$.

The inner product on any Euclidean space is an example of a symmetric positive definite bilinear form.

Riemannian Manifolds

A Riemannian manifold $(\mathcal{M}, g)$ is a differentiable manifold in which each tangent space has an inner product g metric inducing a norm for the tangent vectors.

It is possible to define different metrics on the same manifold to obtain different Riemannian manifolds.

Riemannian Manifolds

A geodesic is a smooth curve that locally joins their points along the shortest path.

The length of the geodesic is defined to be the Riemannian distance between the two points.

Riemannian Manifolds

The exponential map, $\exp_X : T_X \mapsto \mathcal{M}$, maps the vector y in the tangent space to the point on the manifold reached by the geodesic after unit time $\exp_X(y) = 1$.

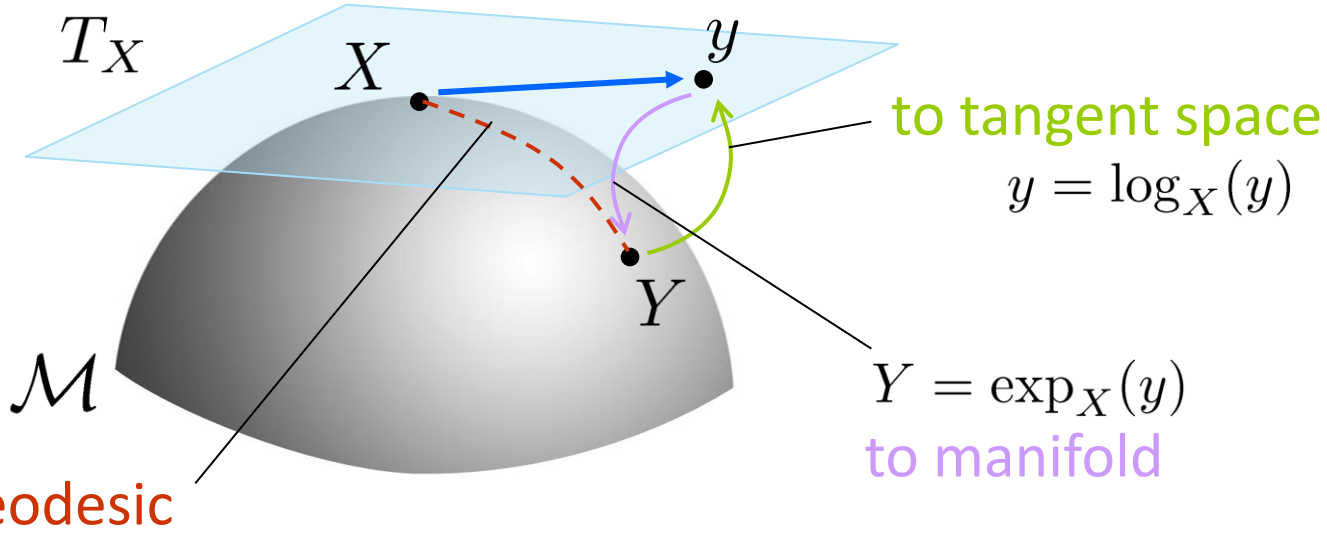
Under the exponential map, the image of the zero tangent vector is the point itself $\exp_X(0) = X$.

Since the velocity along the geodesic is constant, the length of the geodesic is given by the norm of the initial velocity $d(X, \exp_X(y)) = \|y\|_X$

Riemannian Manifolds

For each point on the manifold, the exponential map is a diffeomorphism (one-to-one, onto, differentiable mapping in both directions).

The inverse mapping $\log_X : \mathcal{M} \mapsto T_X$ is uniquely defined only around the neighborhood of the point X.



Riemannian Manifolds

From the definition of geodesic and the exponential map, the distance between the points on manifold can be computed by

$$\begin{aligned} d(X,Y) &= d(X, \exp_X(y)) \\ &= < \log_X(Y), \log_X(Y) >_X \\ &= \| \log_X(Y) \|_X \\ &= \|y\|_X \end{aligned}$$

For Riemannian manifolds endowing an inverse mapping

$$d(X,Y) = \| \log_X(X^{-1}Y) \|$$

Riemannian Manifolds

Lie groups are well known examples of Riemannian manifolds with the structure of an analytic manifold, i.e. multiplication and inversion are smooth maps.

The most frequently occurring Lie groups are sets of matrices, i.e., each element in the group is a matrix and the group operation is matrix multiplication.

Special orthogonal group: rotations in 3D

Euclidean motion group: 3D rigid motions

Covariance matrices

Laplacian matrices

Riemannian Manifolds

For symmetric matrices, the ordinary matrix exponential and logarithm operators can be computed easily.

Let $\Sigma = \mathrm{U}\mathrm{D}\mathrm{U}^T$ be the eigenvalue decomposition of a symmetric matrix. The exponential series is

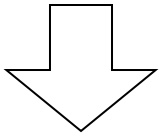
$$\exp(\Sigma) = \sum_{k=0}^{\infty} \frac{\Sigma^k}{k!} = \mathrm{U}\exp(\mathrm{D})\mathrm{U}^T$$

$$\log(\Sigma) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}(\Sigma - \mathrm{I})^k = \mathrm{U}\log(\mathrm{D})\mathrm{U}^T$$

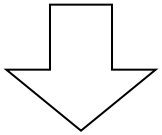
Riemannian Manifolds

Manifold operations for Lie groups are

$$\exp_X(y) = X \exp(X^{-1}y)$$
$$\log_X(Y) = X \log(X_{-1}Y)$$



$$\exp_X(y) = X^{\frac{1}{2}} \exp\left(X^{-\frac{1}{2}} y X^{-\frac{1}{2}}\right) X^{\frac{1}{2}}$$
$$\log_X(Y) = X^{\frac{1}{2}} \log\left(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}}\right) X^{\frac{1}{2}}$$



$$d(X,Y) = \| \log(X^{-1}Y) \|.$$

Riemannian Manifolds

The distance between two points is measured

$$\begin{aligned} d^2(X,Y) &= <\log_X(Y), \log_X(Y)>_X \\ &= \mathrm{tr}\left(\log^2(X^{-\frac{1}{2}}YX^{-\frac{1}{2}})\right) \end{aligned}$$

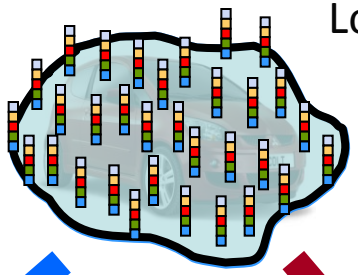
An equivalent form of in terms of joint eigenvalues of X and Y as

$$d(X,Y)^2 = \sum (ln\lambda_k(X,Y))^2$$

X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. IJCV, 2006

Region Covariance

Given image



Low-level features

$$\mathbf{f}(x,y) = [x \quad y \quad I(x,y) \quad I_x(x,y) \quad \dots \quad]_k$$

2nd order statistics:

$$\mathbf{C} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{f}_n - \mu)(\mathbf{f}_n - \mu)^T$$

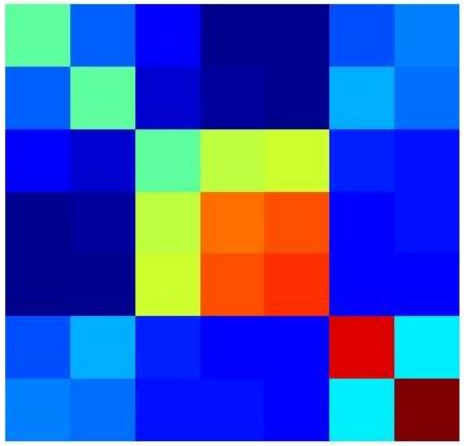
1st order statistics:

Distribution function (histograms)

- Blind to spatial pattern
- Requires many samples to populate
- Noise sensitive

Same histogram

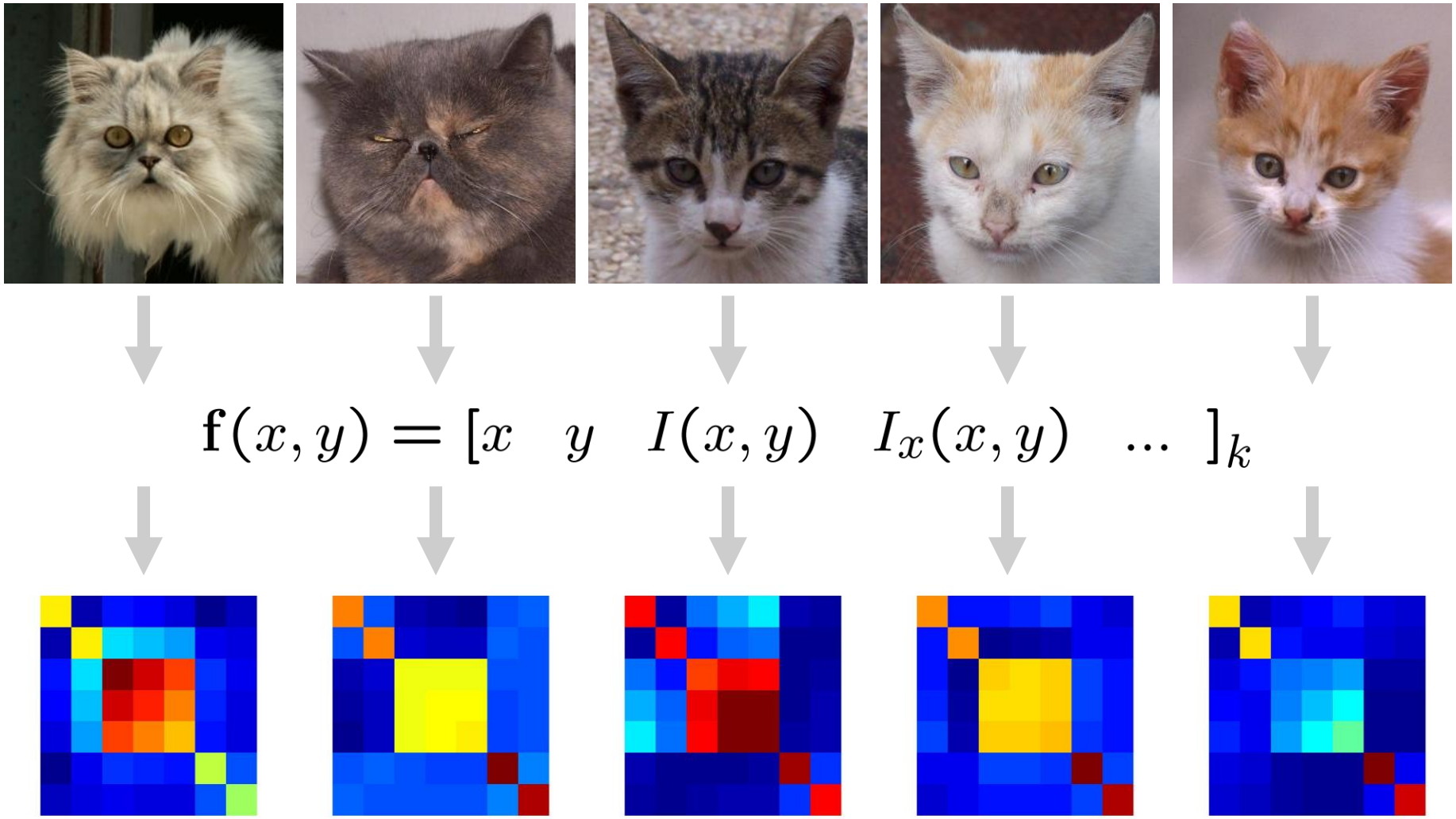
$x \quad y \quad R \quad G \quad B \quad I_x \quad I_y$



Region Covariance

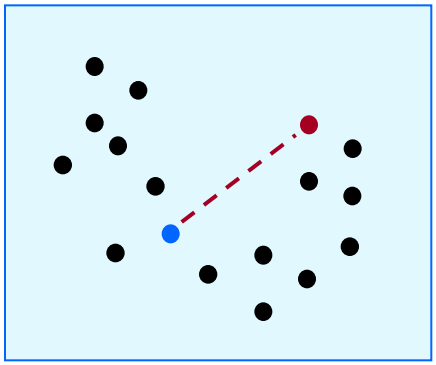
- Symmetric, positive, semi-definite → Lie group
- Low-dimensional: 7 features: 28 coefficients
  - 16 bins color histogram: 4096 coefficients
- Natural way of fusing features
- Invariant to rotation & scale changes, affine transformations, illumination variances

Region Covariance

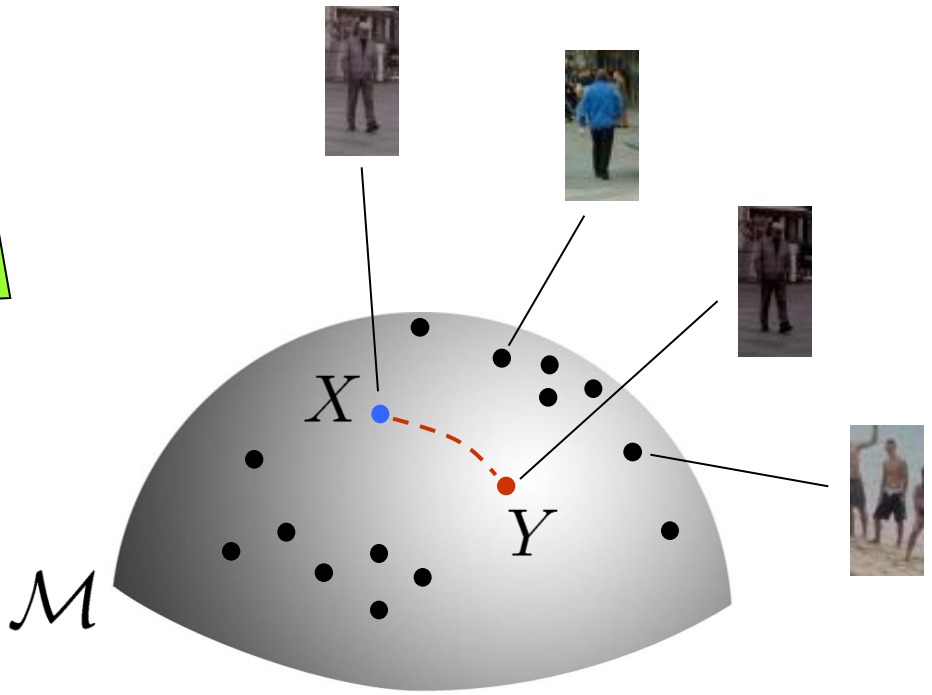


$$\mathbf{f}(x,y) = [x \quad y \quad I(x,y) \quad I_x(x,y) \quad \dots \quad]_k$$

Search on Manifold

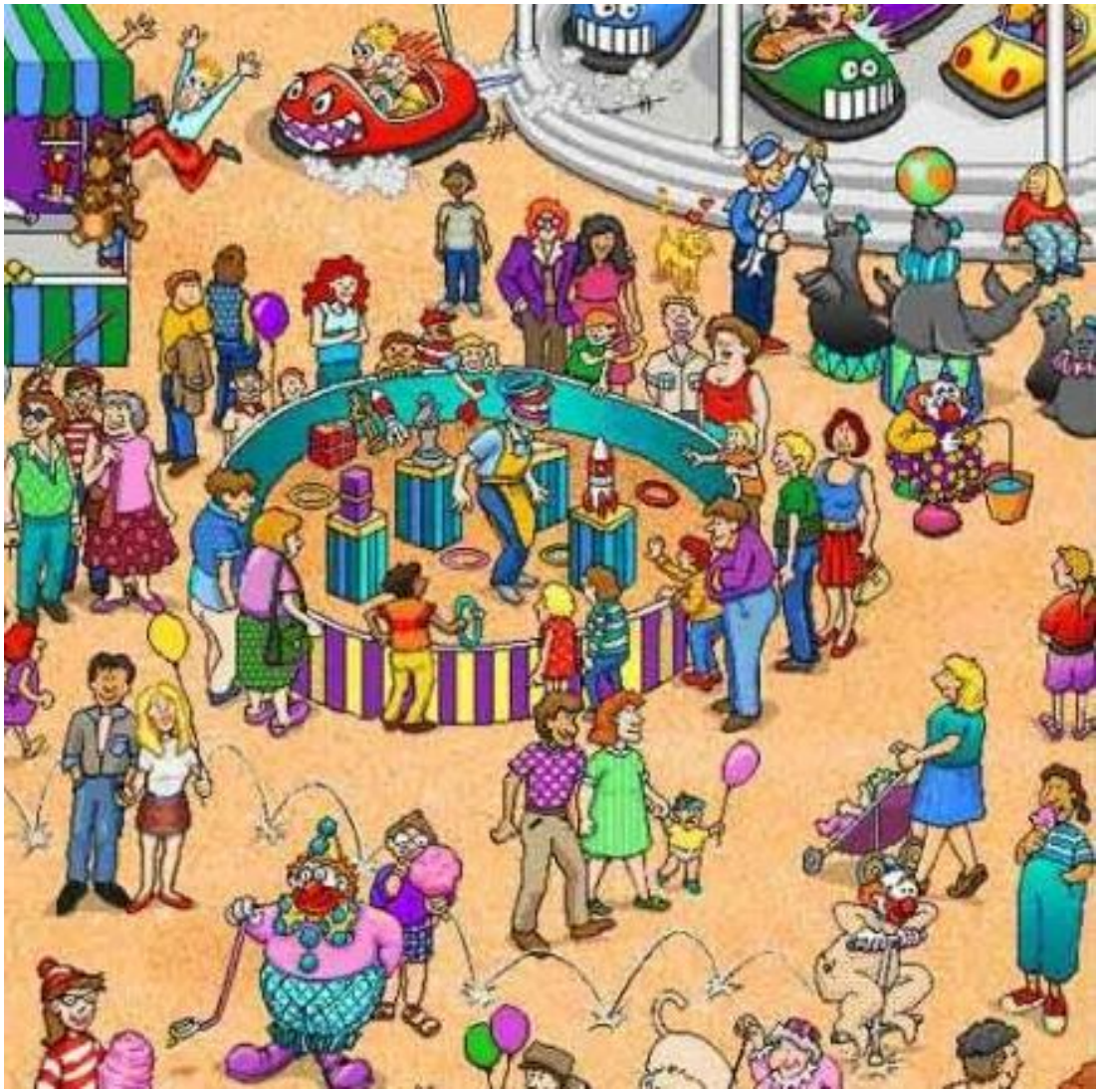


$\mathbb{R}^m$



Search on Manifold

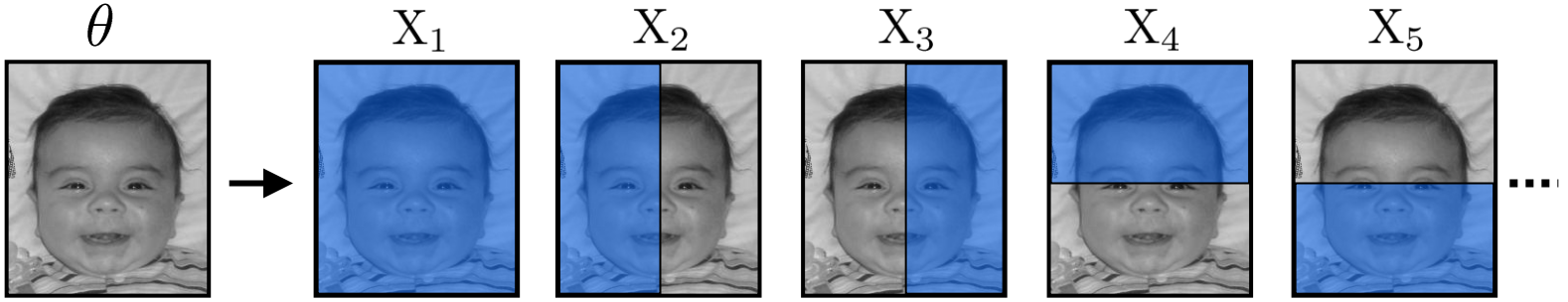
Given



find

Search on Manifold

- Construct multiple covariance matrices from overlapping regions



$$\mathbf{f}(x,y) = [x \quad y \quad R \quad G \quad B \quad I_x \quad I_y \quad I_{xx} \quad I_{yy} \quad]_9$$

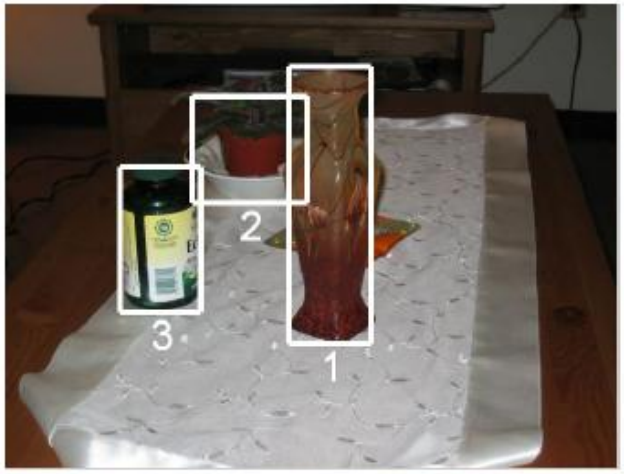
- Search given image for a region having similar covariance matrix
  - 9 different scales (4 smaller, 4 larger, 15% scaling factor between scales)

- Keep the best matching 1000 locations, second phase, repeat the search using all matrices

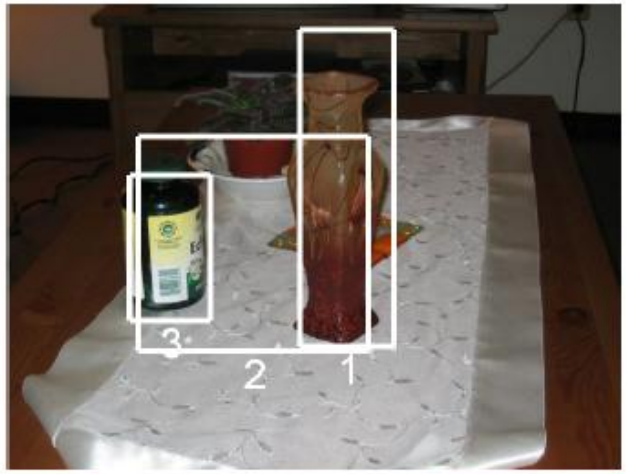
$$d(I,T) = \min_j \left[\sum_{i=1}^{M} d(\mathbf{C}_i^I, \mathbf{C}_i^T) - d(\mathbf{C}_j^I, \mathbf{C}_j^T) \right]$$

Search on Manifold

Given



Manifold



Histogram features

Search on Manifold

Given

Manifold

Histogram features

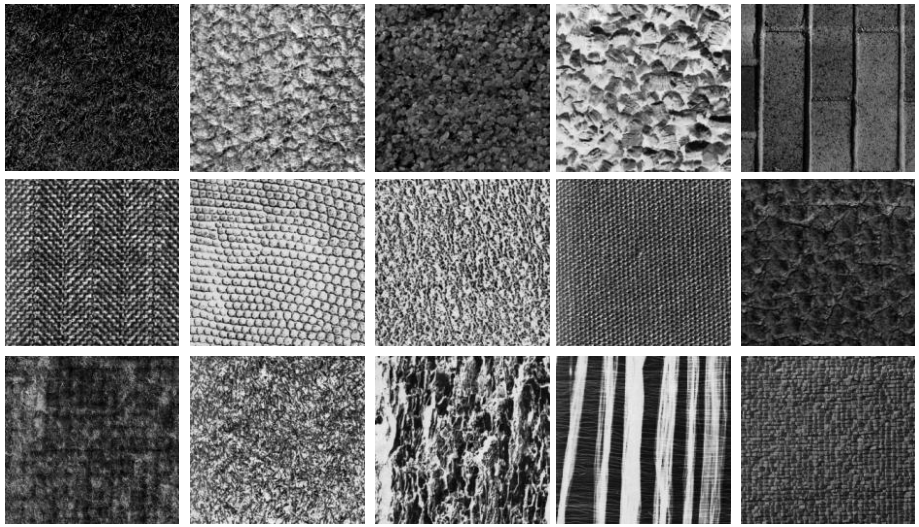
Search on Manifold

Given

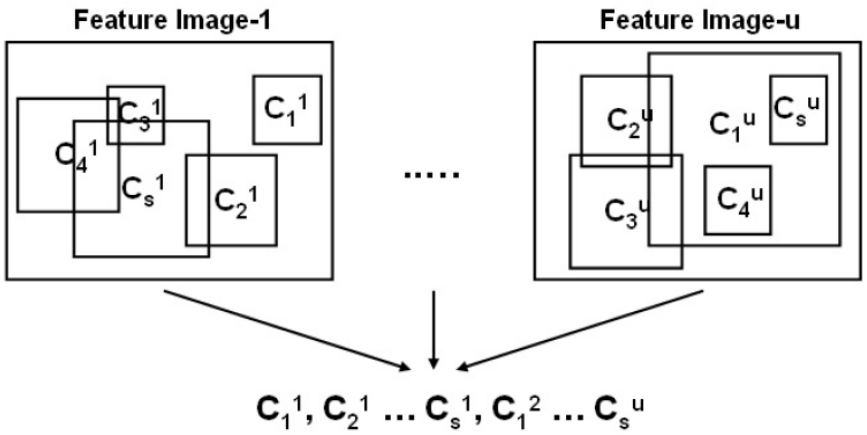
Manifold

Histogram features

K-NN on Manifold



112 Classes



s = 100, k = 5

1% of the computation

	M4	M8	S	LM	Random Covariance
Performance	85.71	94.64	93.30	97.32	97.77

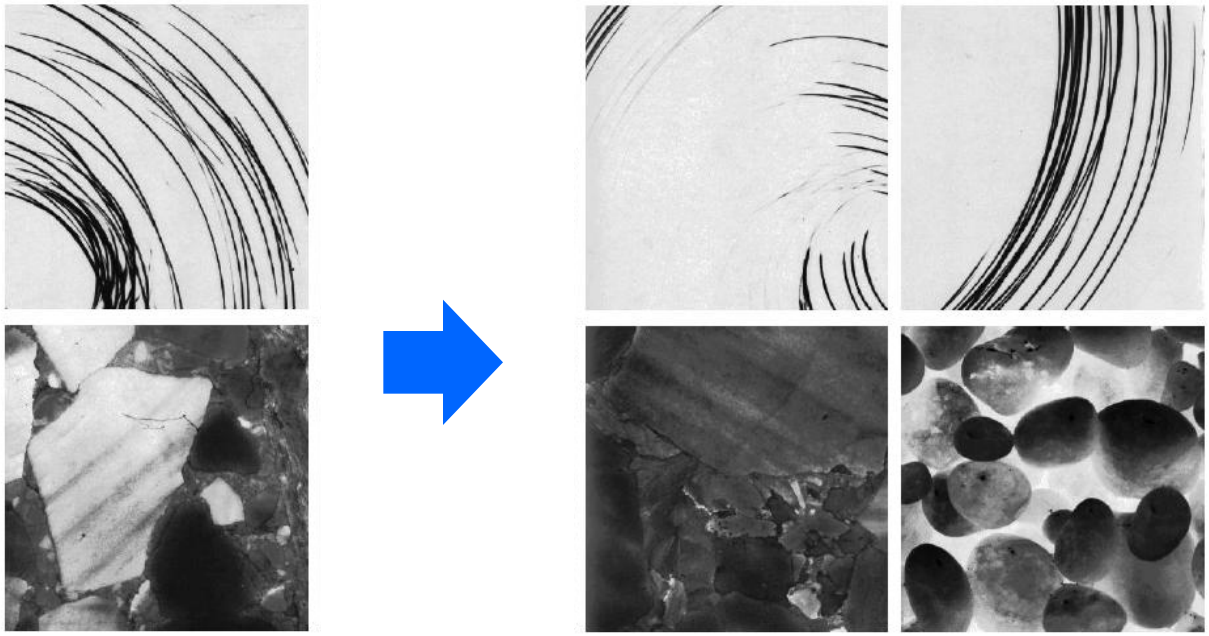
LM: *A combination of 48 anisotropic and isotropic filters (Leung and Malik). The feature space is 48 dimensional (heavy computation)*

S: *A set of 13 circular symmetric filters (Schmid). The feature space is 13 dimensional (heavy computation)*

M4, M8: *Textons (Varma and Zissermann). The texton feature space is 4 and 8 dimensional respectively (very heavy computation)*

K-NN on Manifold

3 misclassified samples (112 classes, 109 correct / 3 false)

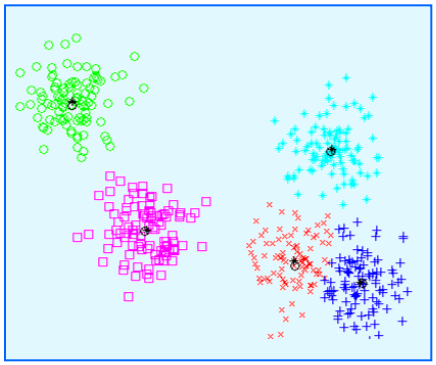


Test samples

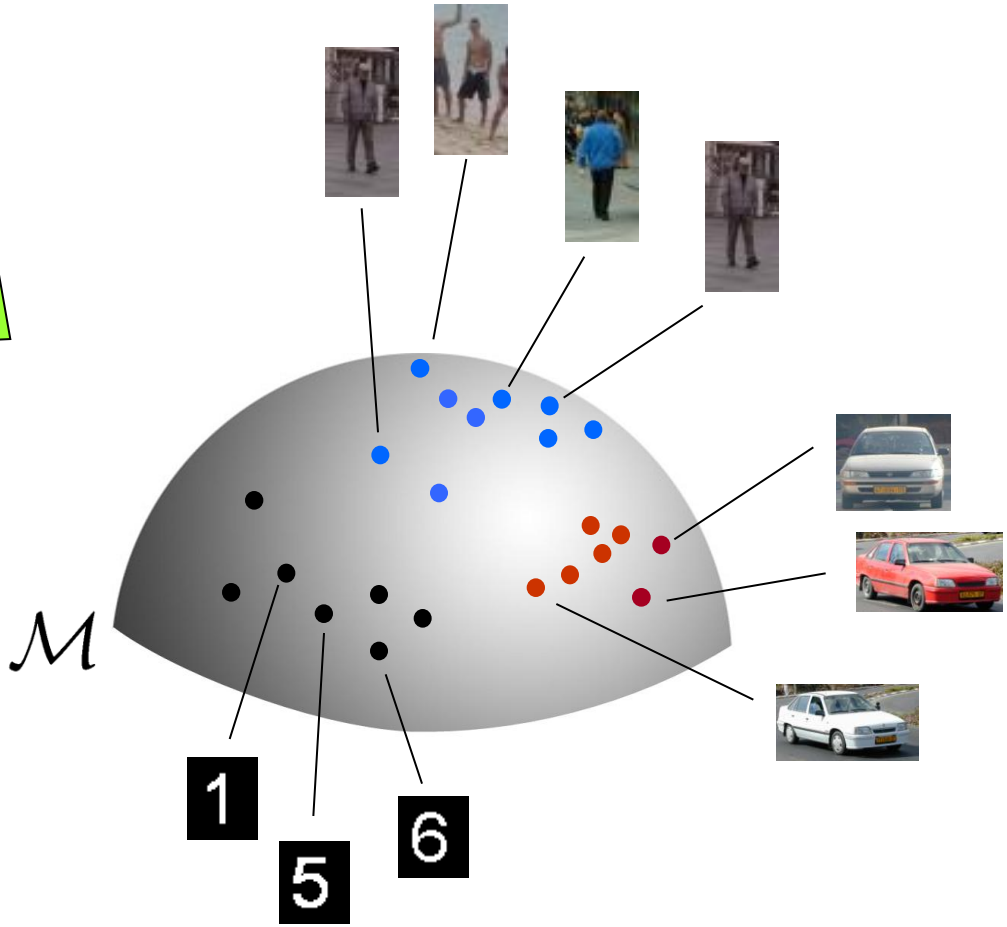
Training samples from the same class

Training samples from the predicted class (however, these matches are counted as misclassification)

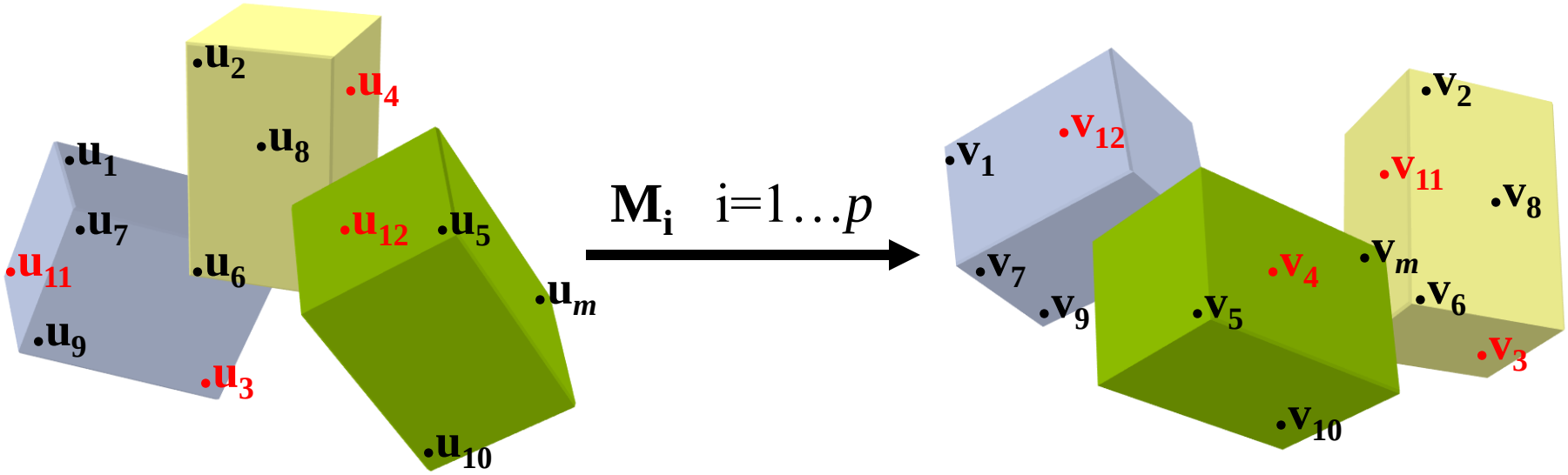
Clustering on Manifold



$\mathbb{R}^m$



Multiple 3D Motion Estimation



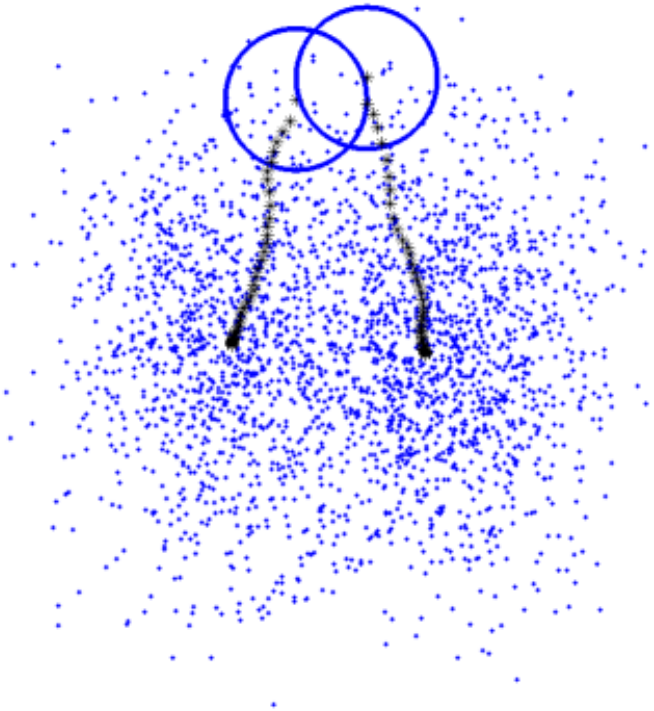
inliers outliers

- Estimate the number of motions p
- Estimate the motion parameters $\mathbf{M}_i \quad i=1\ldots p$

- Existing approaches are only iterative solutions such as RANSAC

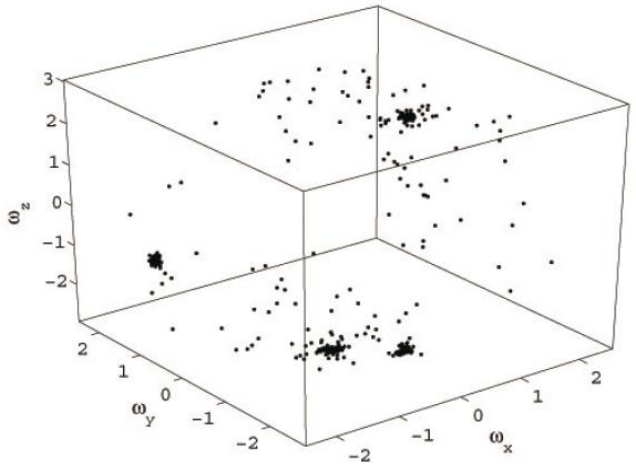
Mean-Shift (Mode Seeking)

- Start at each point
  - Estimate the density mean within a kernel
  - Shift the kernel
  - Iterate
- Combine modes

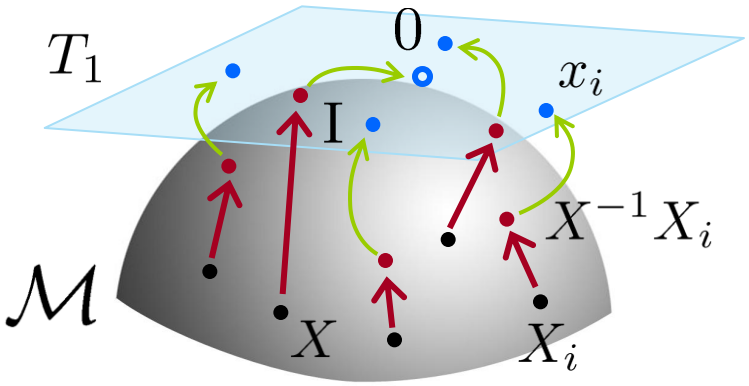


Hypotheses Generation

- Any three point correspondences generate a motion hypothesis
- Generate m hypotheses
- m hypotheses are either samples from the multiple motion distribution or random noise (outliers for the distribution)
- Number of significant modes are the number of motion groups p, and the modes are the motion parameters $\mathbf{M}_i$



Mean-Shift on Manifold



Iteration 1

$$x_i = \log(X^{-1}X_i)$$

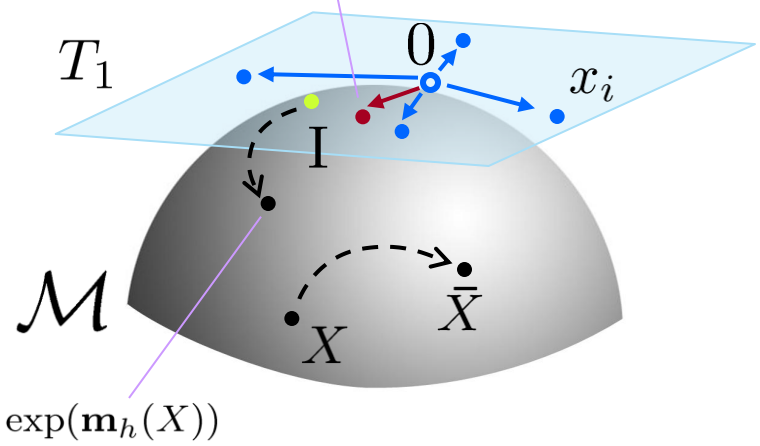
Left multiplying by inverse we transform the neighborhood of X to neighborhood of $\mathbf{I}$

Points on the manifold are mapped to tangent space

X is mapped to 0

Mean-Shift on Manifold

$$\mathbf{m}_h(X) = \frac{\sum_{i=1}^{n} x_i g\left(\left\|\frac{x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{x_i}{h}\right\|^2\right)}$$



Iteration 1

Mean-shift vector is the average of the points weighted by derivative of Gaussian

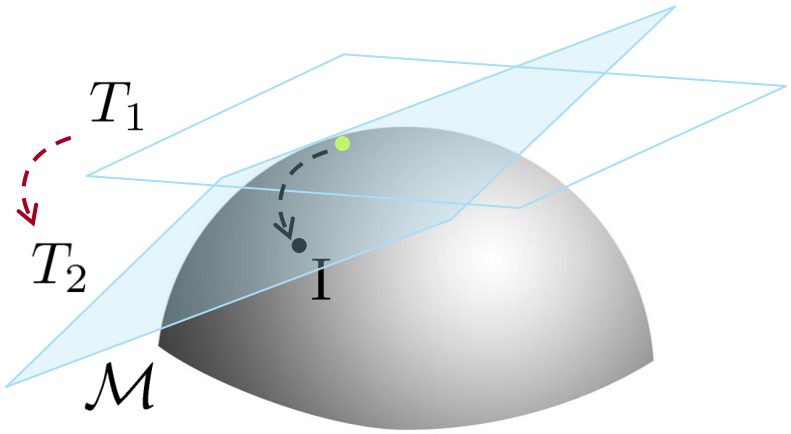
Computed average is a first order approximation to the true mean

Take the exponential of mean-shift vector and map it onto manifold.

$$\bar{X} = X \exp(\mathbf{m}_h(X))$$

Right multiplication of vector with X updates the location of X

Mean-Shift on Manifold



Iteration 2

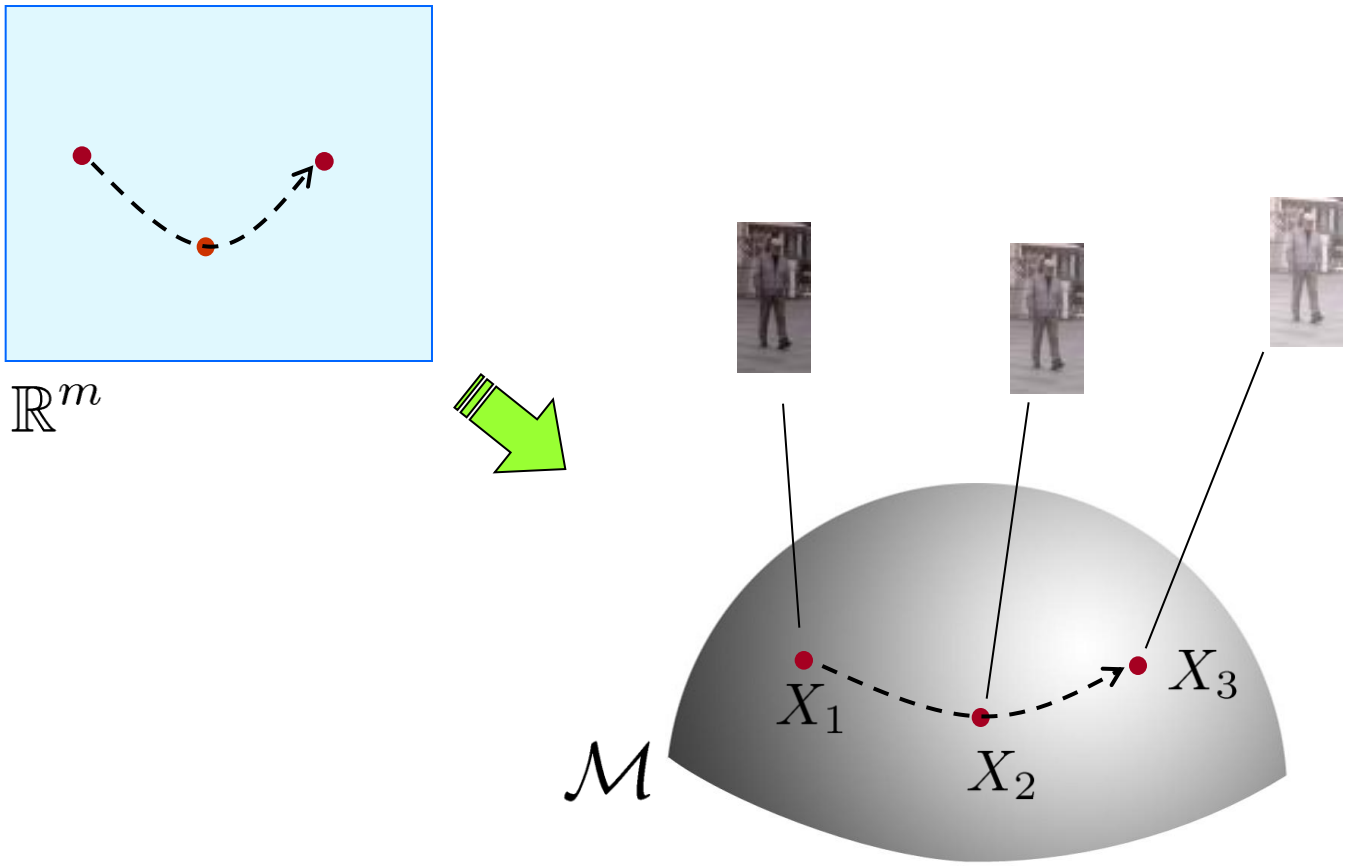
Iterate the process until convergence.

Usually 4~5 iterations suffices.

Mean Shift on Lie Groups

Less than half of the 83 points could be matched correctly

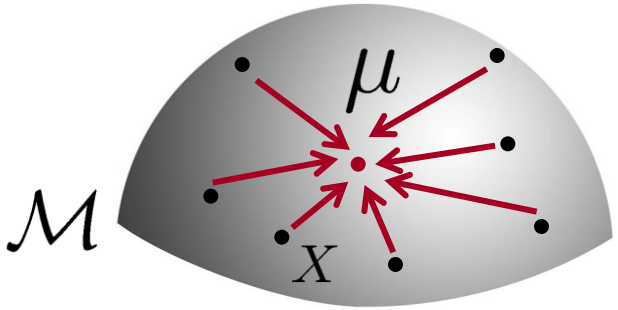
Model Update on Manifold



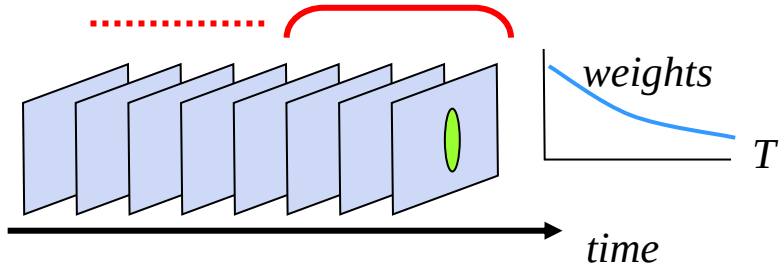
Model Update on Manifold

- Mean of the points on the manifold is the point that minimizes sum of squared geodesic distances

$$\mu = \arg\min_{Y \in M} \sum_{i=1}^{N} d^2(X_i, Y)$$



- Differentiating the error function wrt $\mathbf{Y}$, gives the following gradient descent



$$\mu^{t+1} = \exp_{\mu^t}\left[\frac{1}{N}\sum_{i=1}^{N} \log_{\mu^t}(X_i)\right]$$

$$\mu^{t+1} = \exp_{\mu^t}\frac{1}{d^*}\left[\frac{1}{N}\sum_{i}^{N} d^{-1}(X_i, X^*)\log_{\mu^t}(X_i)\right]$$

Model Update on Manifold

No model update;
Detection rate is 47.7%.

Model update
Detection rate is 100%.

Model Update on Manifold

Model Update on Manifold

Model Update on Manifold

Model Update on Manifold

Covariance

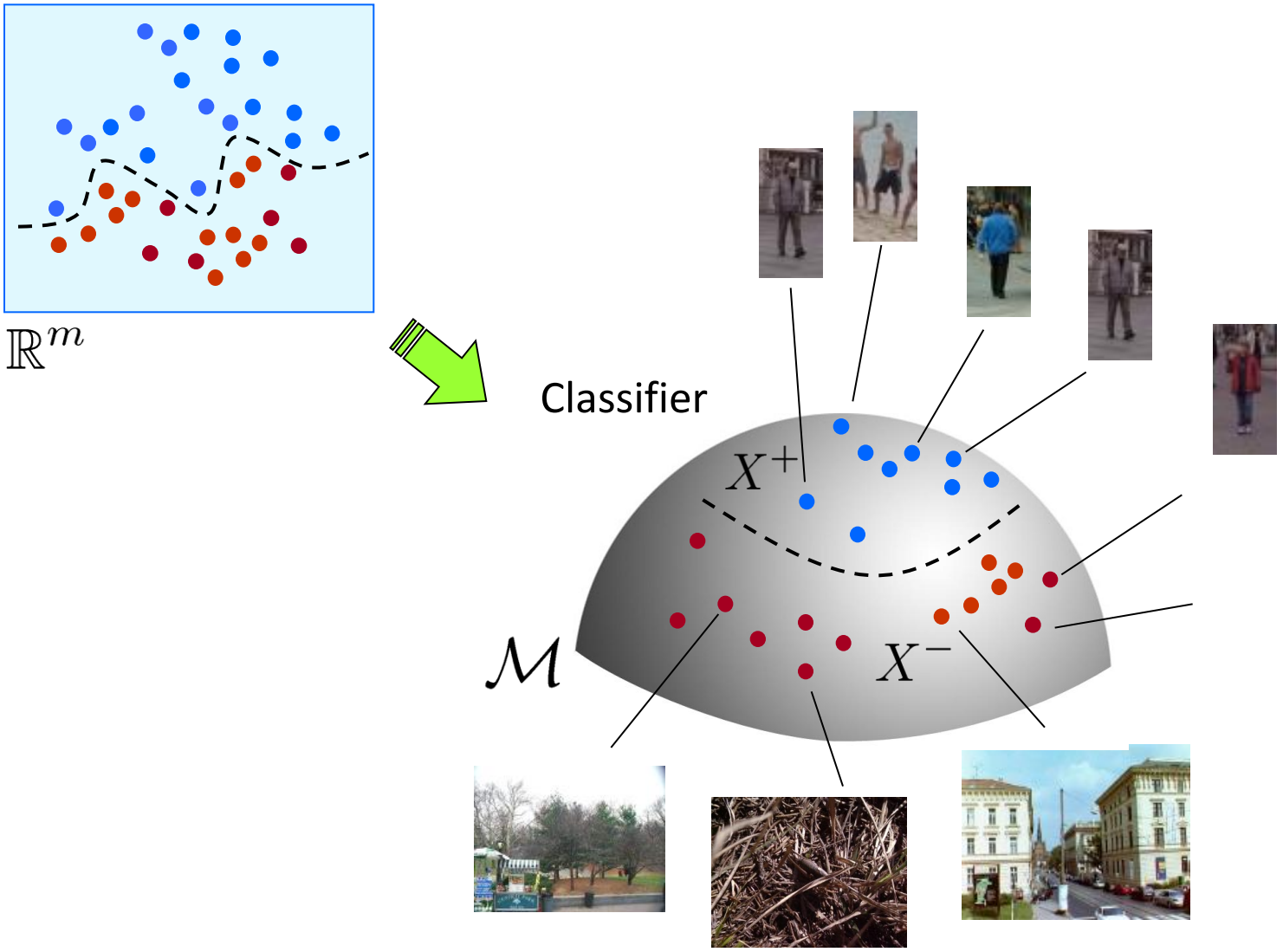
	RGB	HS-only
Covariance tracking	95.6%	93.3%
Histogram matching	48.7%	64.0%

Model Update on Manifold

Covariance

	$\sigma_\eta^2 = 0$	$= 0.01$	$= 0.1$	$= 0.3$
Covariance tr.	97.4%	94.3%	89.0%	70.6%
Histogram mat.	72.8%	65.2%	42.6%	18.9%

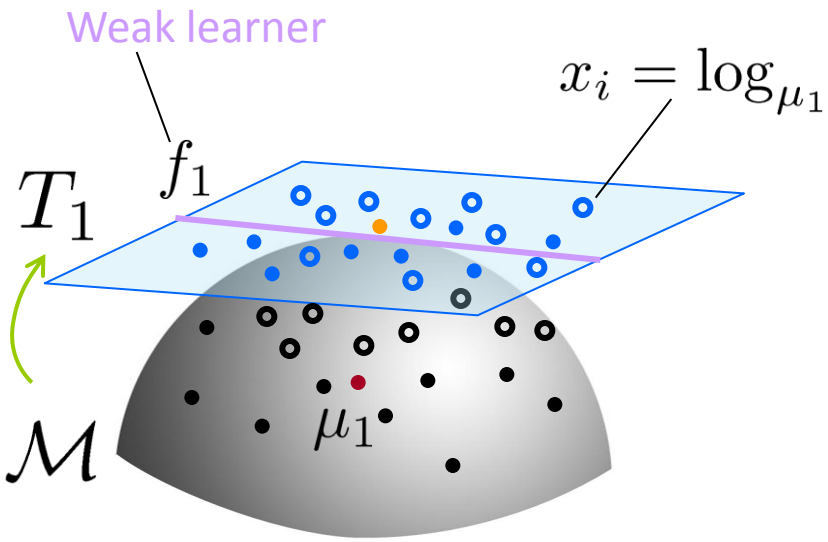
Classifiers on Manifold



Classifiers on Manifold

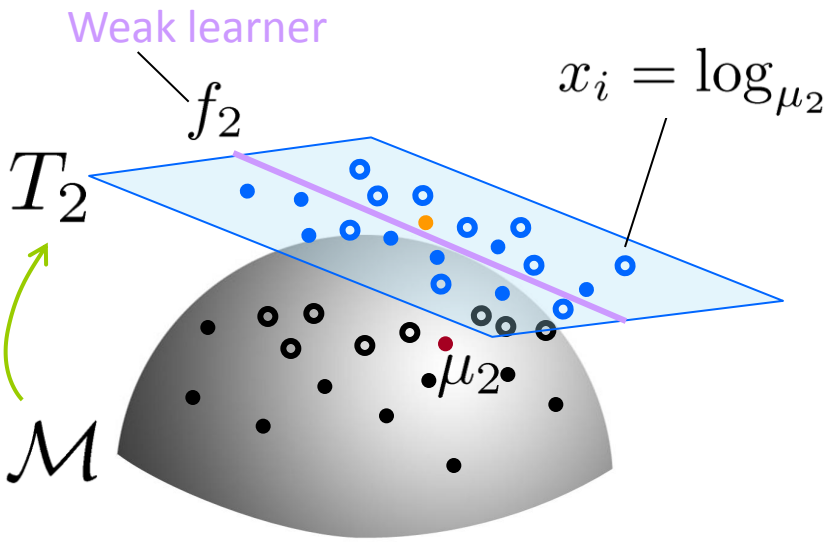
- At each iteration, we compute the weighted mean of the points where the weights are adjusted through boosting

- Map the points to the tangent space at the mean and learn a weak classifier on this vector space



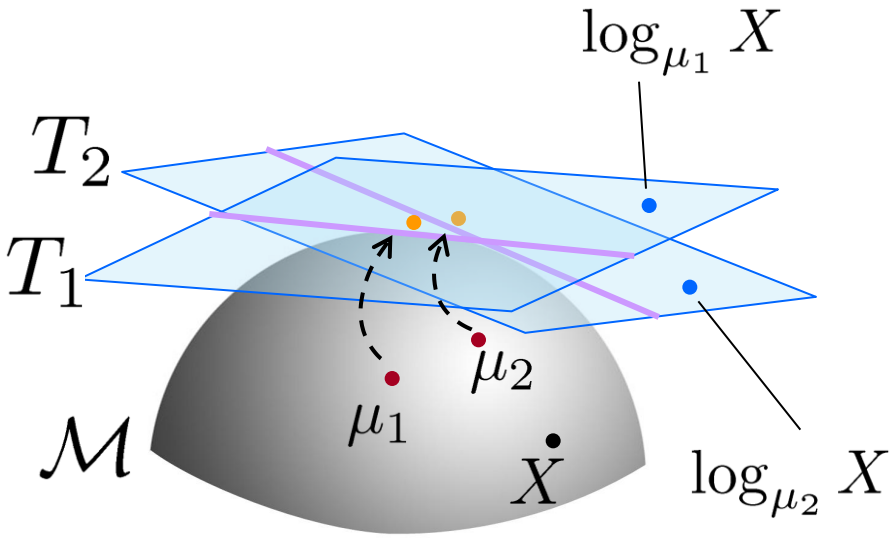
Classifiers on Manifold

- Weights of the samples which are misclassified increase
- Weighted mean moves towards these points producing more accurate classifier
- This approach minimizes the approximation error through averaging over several weak classifiers



Boosting on Manifold

- Same transformations are applied to the test sample and the weak learners are evaluated

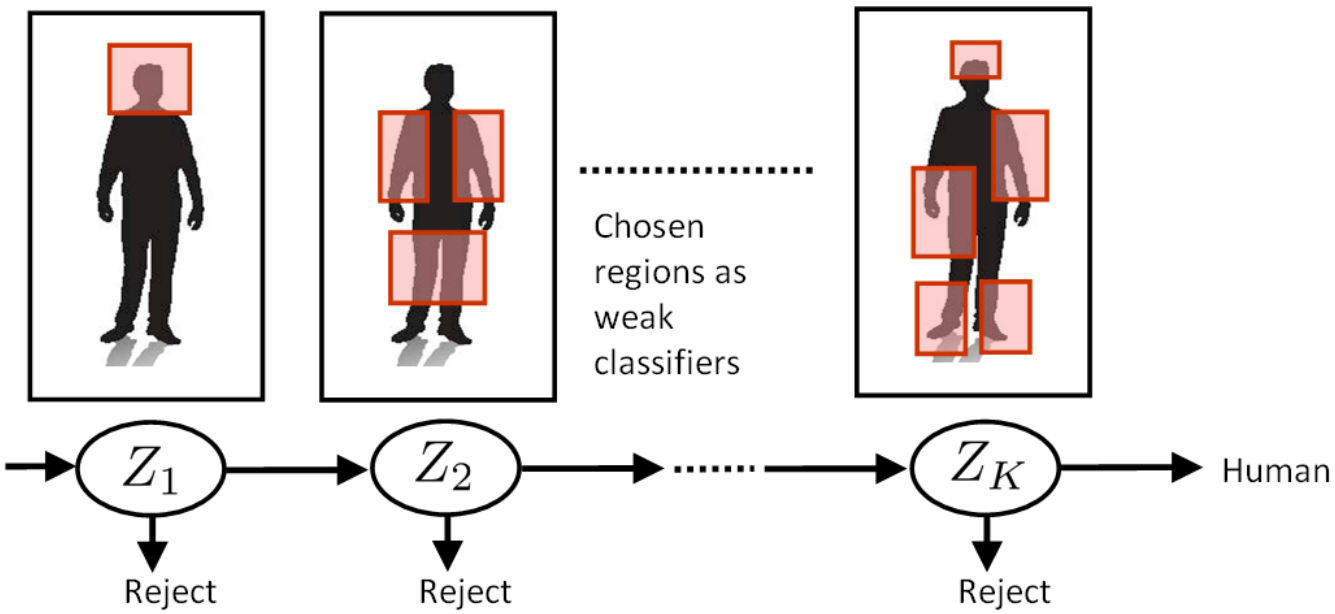


Boosting on Manifold

Input: Training set $\{(\mathbf{X}_i, y_i)\}_{i=1...N}$, $\mathbf{X}_i \in \mathcal{M}$, $y_i \in \{0, 1\}$

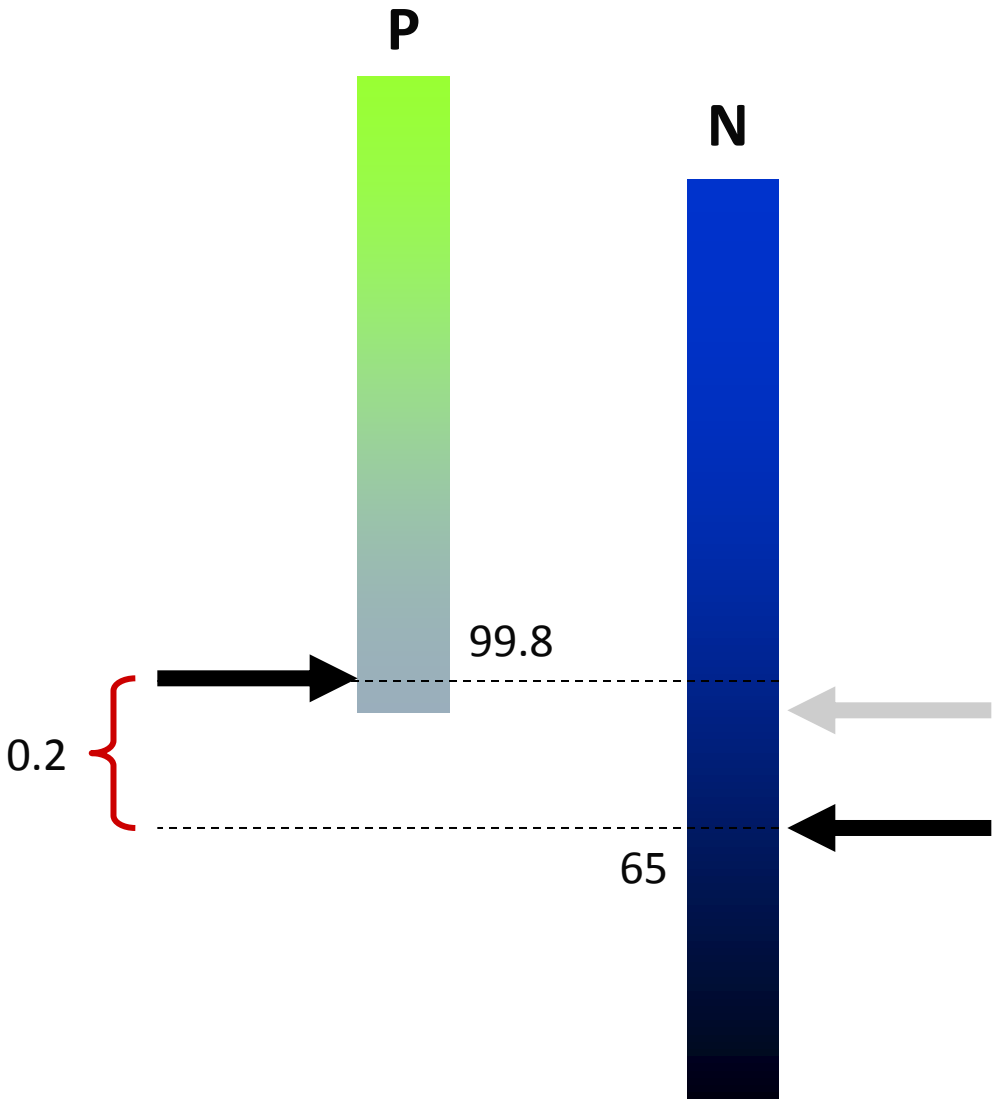
- Start with weights $w_i = 1/N$, $i = 1...N$,
  $F(\mathbf{X}) = 0$ and $p(\mathbf{X}_i) = \frac{1}{2}$
- Repeat for $l = 1...L$
  - Compute the response values and weights
    $z_i = \frac{y_i - p(\mathbf{X}_i)}{p(\mathbf{X}_i)(1-p(\mathbf{X}_i))}$
    $w_i = p(\mathbf{X}_i)(1 - p(\mathbf{X}_i))$.
  - Compute weighted mean of the points
    $\boldsymbol{\mu}_l = \arg\min_{\mathbf{Y} \in \mathcal{M}} \sum_{i=1}^{N} w_i d^2(\mathbf{X}_i, \mathbf{Y})$. $(*)$
  - Map the data points to the tangent space at $\boldsymbol{\mu}_l$
    $\mathbf{x}_i = \log_{\boldsymbol{\mu}_l}(\mathbf{X}_i)$. $(*)$
  - Fit the function $g_l(\mathbf{x})$ by weighted least-square regression of z_i to $\mathbf{x}_i$ using weights w_i.
  - Update $F(\mathbf{X}) \leftarrow F(\mathbf{X}) + \frac{1}{2} f_l(\mathbf{X})$ and
    $p(\mathbf{X}) \leftarrow \frac{e^{F(\mathbf{X})}}{e^{F(\mathbf{X})} + e^{-F(\mathbf{X})}}$.
- Output the classifier sign
  $[F(\mathbf{X})] = \text{sign}\left[\sum_{l=1}^{L} f_l(\mathbf{X})\right]$

Boosting on Manifold

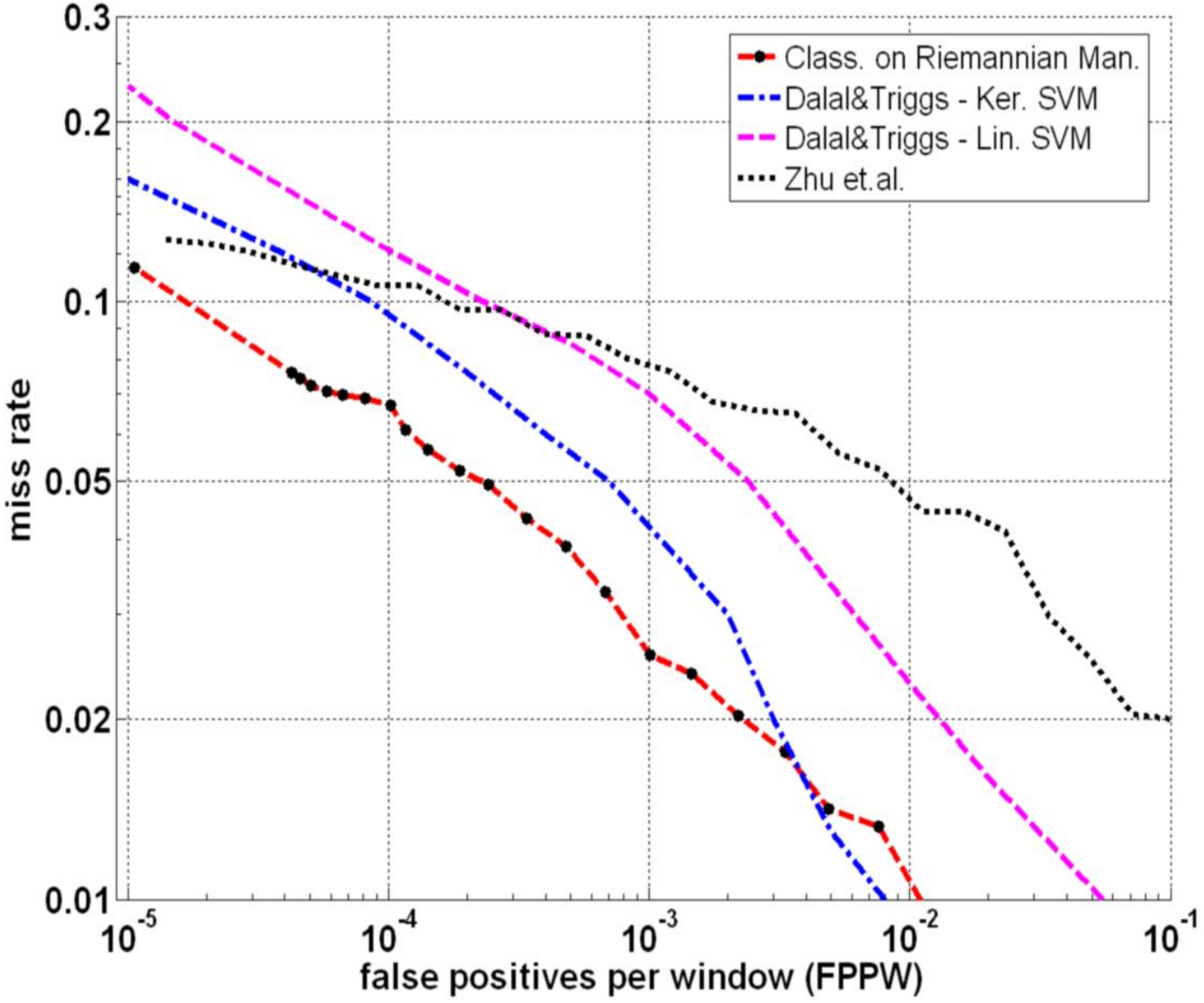


- For human detection, K=30 LogitBoost classifiers is combined with rejection cascade
- Each level of cascade detector is optimized to 99.8% true positive rate, and 65% false positive rate
- Weak classifiers are linear regression functions learned on the tangent space (m = 36 dimensional vector space)
  - Sample subwindows, add the best classifier to the strong classifier

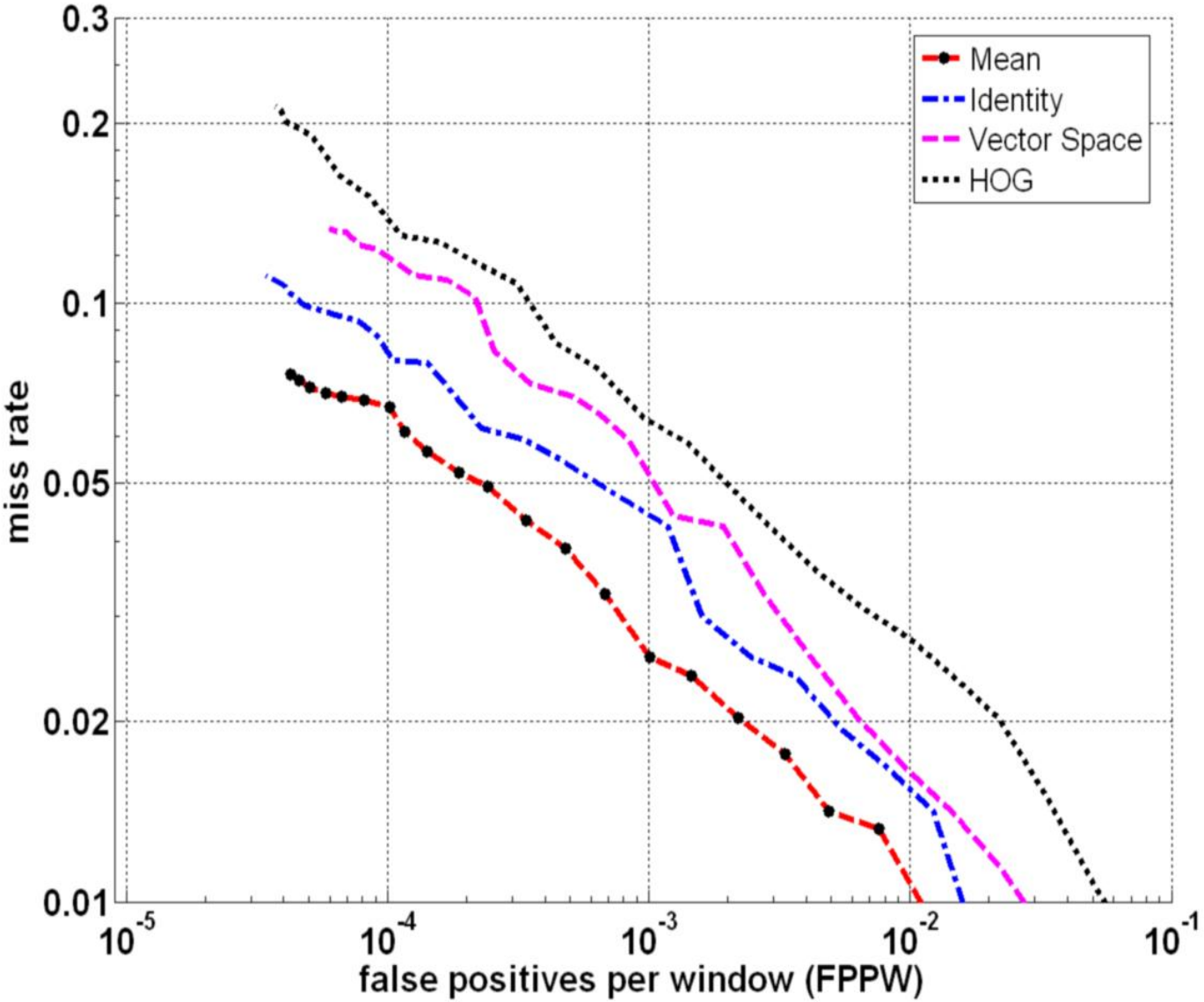
Boosting on Manifold



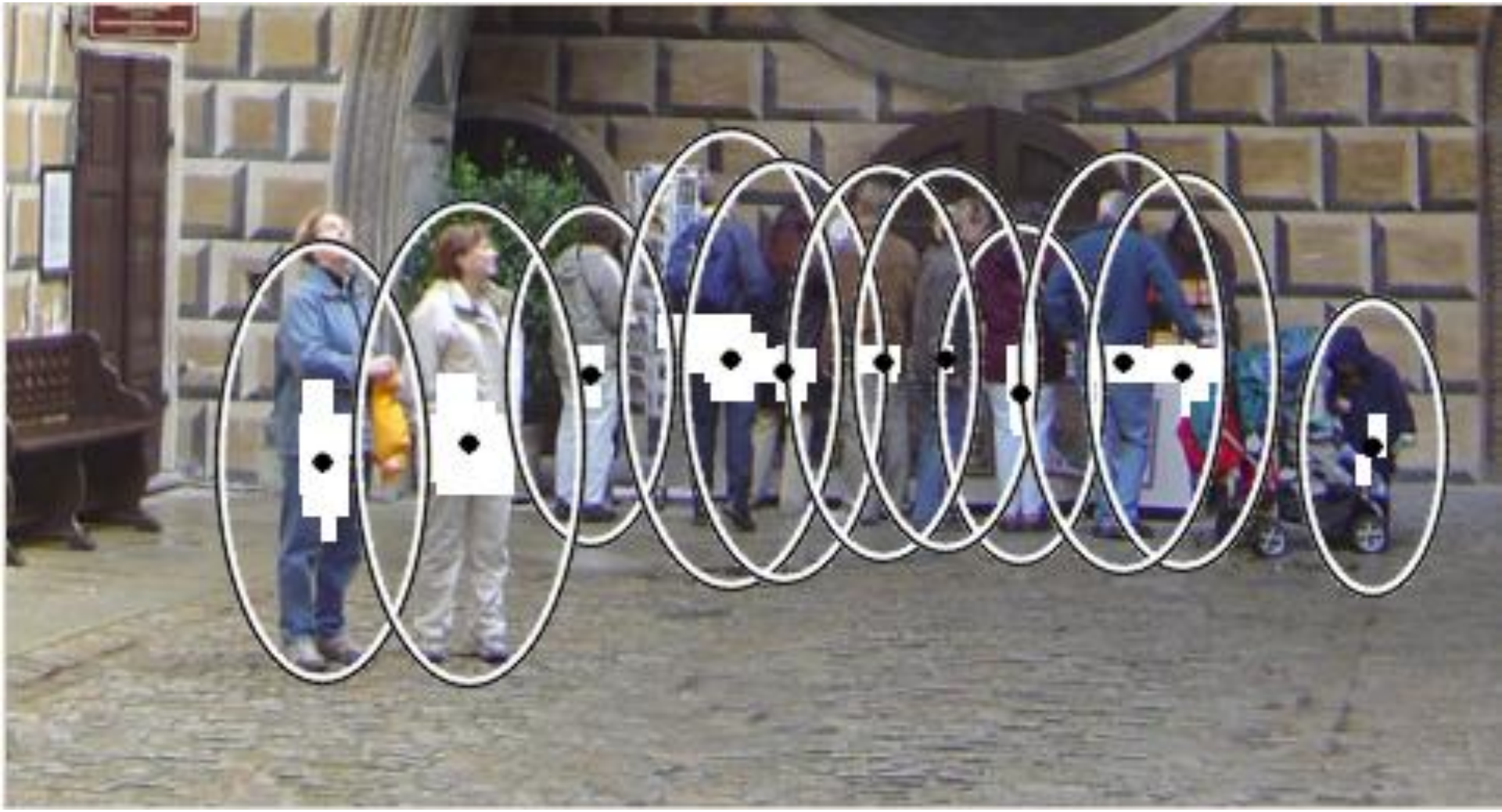
Boosting on Manifold



Boosting on Manifold



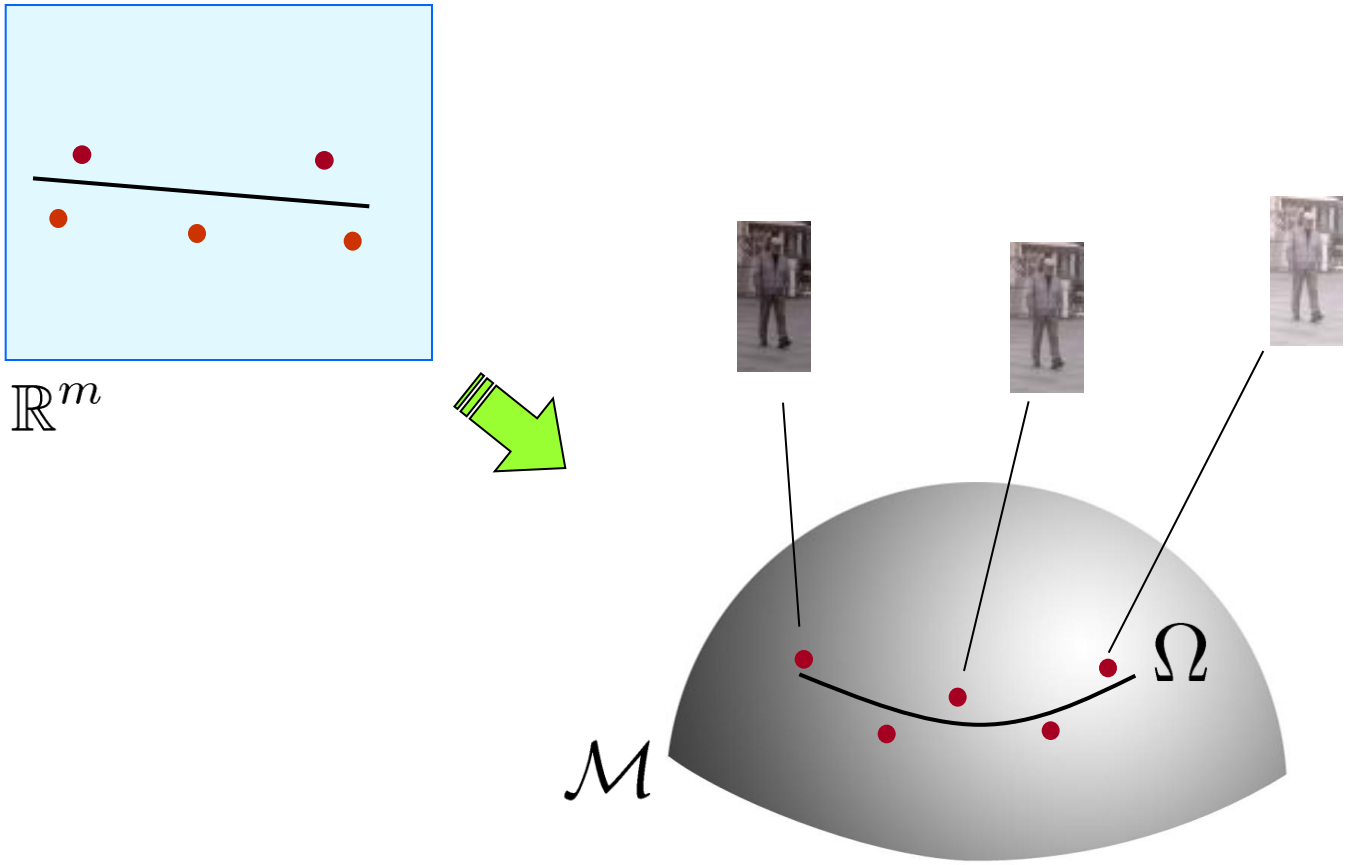
Boosting on Manifold



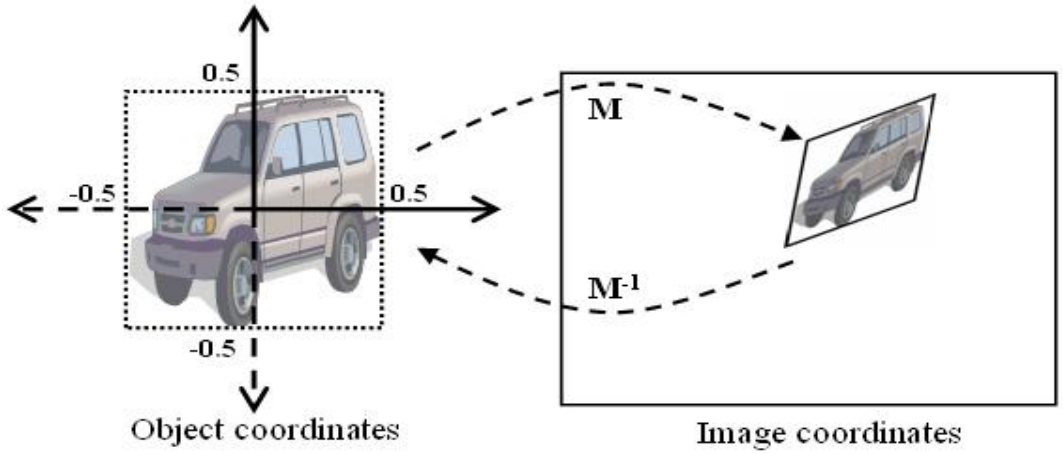
Boosting on Manifold

Boosting on Manifold

Regression on Manifold



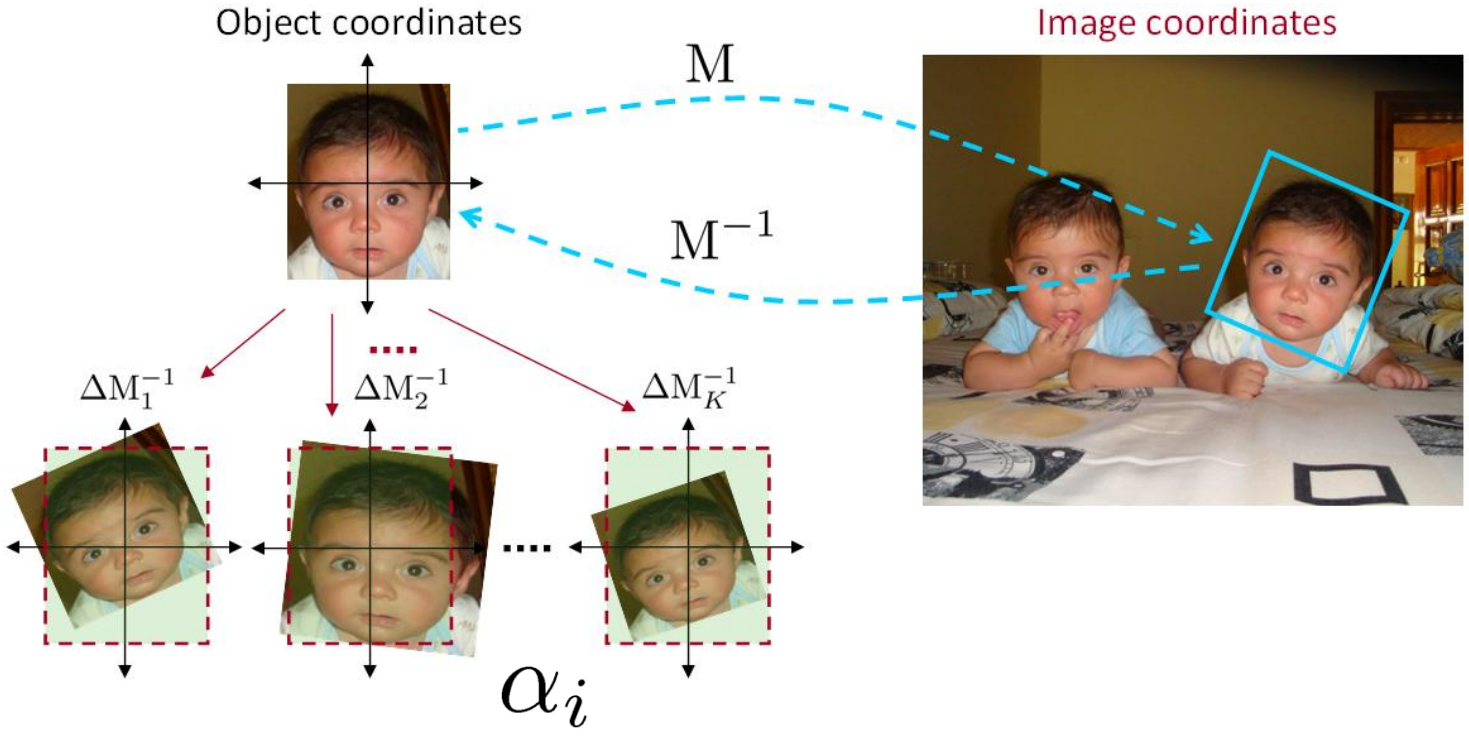
Regression on Manifold



$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ 0 & 1 \end{pmatrix}$$

- The region is represented with several gradient weighted orientation histograms computed at a regular grid

Regression on Manifold



- Model the interframe motion incrementally

$$\mathbf{M}_t = \mathbf{M}_{t-1}.\Delta\mathbf{M}_t$$

where $\Delta\mathbf{M}_t$ corresponds to motion of target

Regression on Manifold

(α_i, X_i) are the pairs of observed data $\alpha \in \mathbb{R}^d$ in vector space and the corresponding points on the manifold.

The regression function φ maps the vector space data onto the manifold

$$\varphi : \mathbb{R}^d \mapsto \mathcal{M}$$

An objective function is defined as the sum of the squared geodesic distances between the estimations $\varphi(\alpha_i)$ and the points X_i

$$J = \sum_i d^2 \left[\varphi(\alpha_i), X_i\right].$$

Regression on Manifold

Lie algebra exists, thus the objective function can be written as

$$J = \sum_i \left\| \log \left[\varphi^{-1}(\alpha_i) X_i \right] \right\|^2 \approx \sum_i \left\| \log \left[\varphi(\alpha_i) \right] - \log \left[X_i \right] \right\|^2$$

up to the first order terms.

The regression function φ can be written as

$$\varphi(\alpha_i) = \exp\left(\alpha_i^T \Omega\right)$$

where the function $\Omega : \mathbb{R}^d \mapsto \mathbb{R}^n$ estimates the tangent vectors $\log(X_i)$ on the Lie algebra.

Regression on Manifold

The objective function becomes

$$J = \sum_i \left\| \alpha_i^T \Omega - \log\left[X_i\right] \right\|^2$$

$$\mathbf{Y} = \Omega \mathbf{X}$$

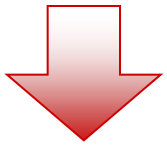
$$\mathbf{X} = \begin{bmatrix} [\alpha_1]^T \\ \vdots \\ [\alpha_k]^T \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} [\log(X_1)]^T \\ \vdots \\ [\log(X_k)]^T \end{bmatrix}$$

$$J = tr[(\mathbf{X}\Omega - \mathbf{Y})^T (\mathbf{X}\Omega - \mathbf{Y})]$$

Regression on Manifold

- To avoid over-fitting, introduce an additional constraint on the size of the regression coefficients (ridge regression). Solution is given by

$$J = tr[(\mathbf{X}\Omega - \mathbf{Y})^T(\mathbf{X}\Omega - \mathbf{Y})] + \lambda\|\Omega\|^2$$

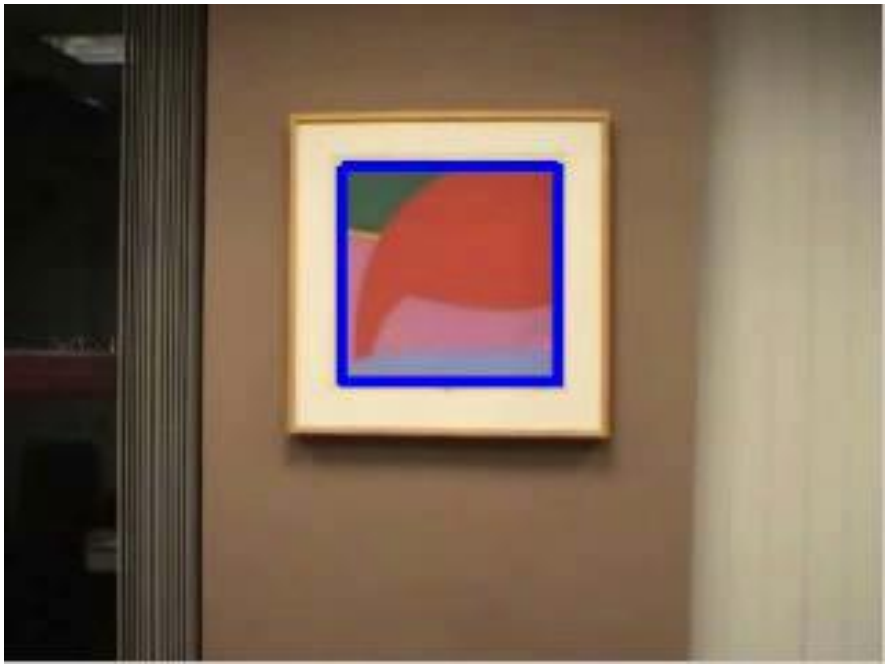


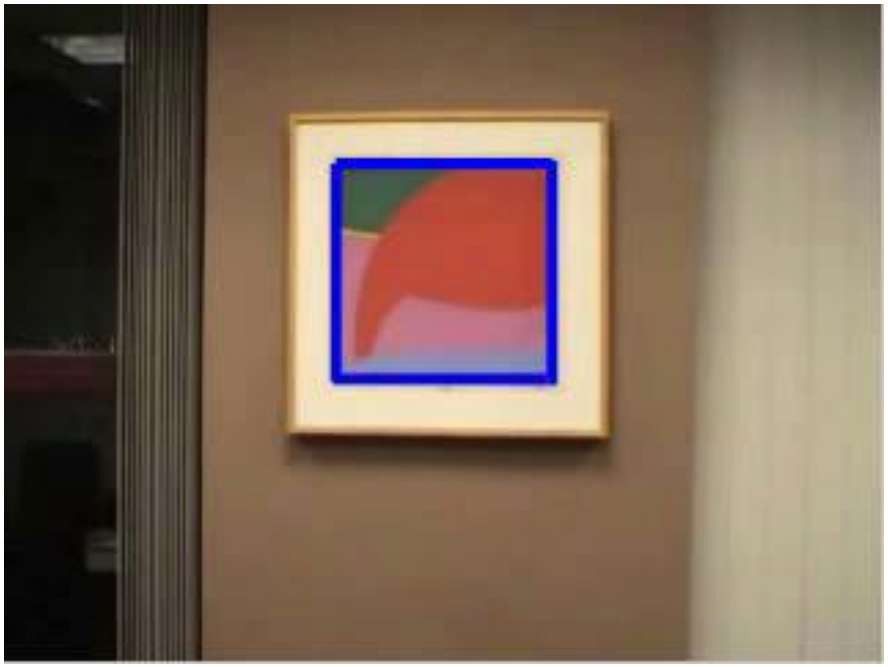
$$\Omega = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{Y}$$

Regression on Manifold

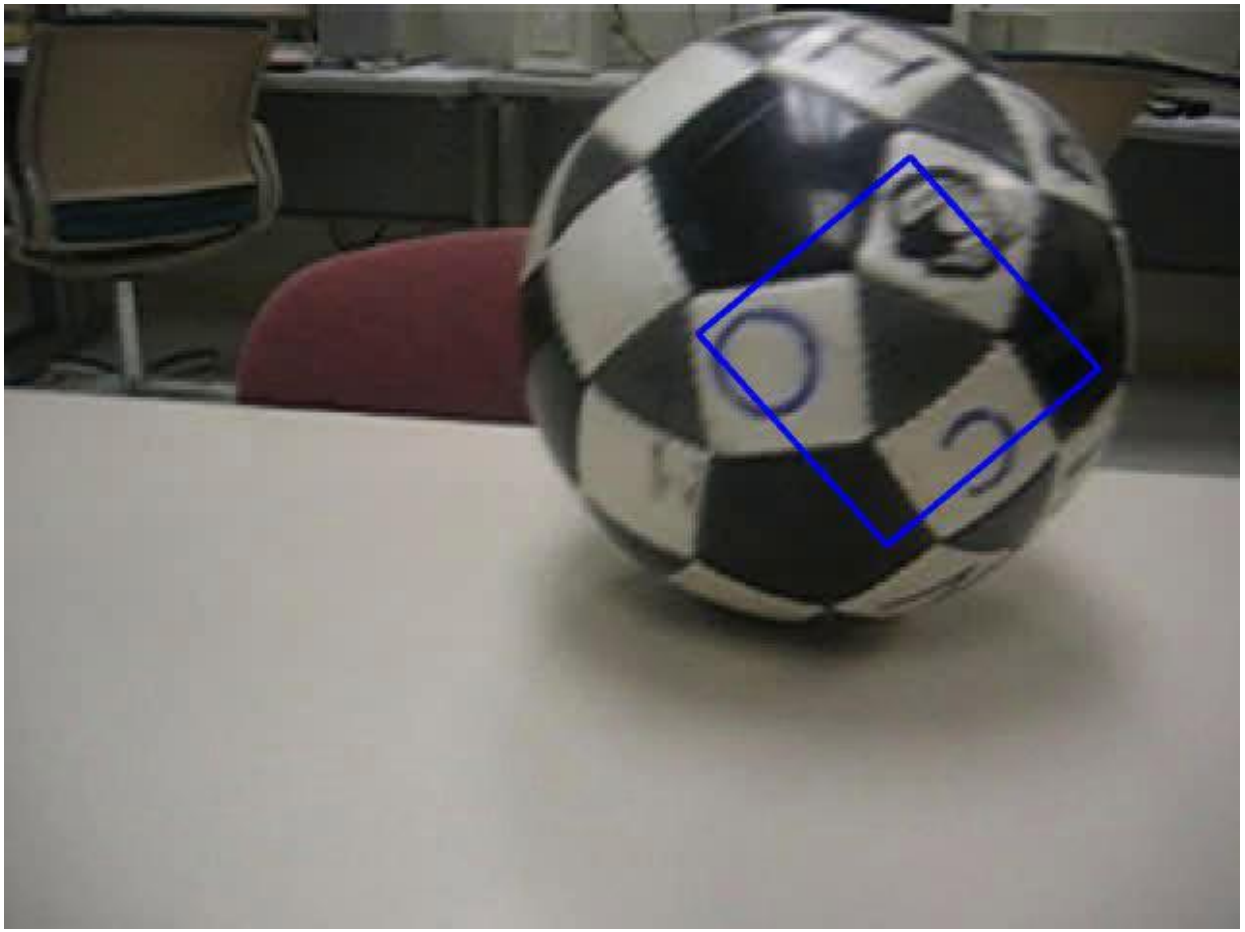
For real time tracking, keep the size of the training set small; n=200

Regression on Manifold

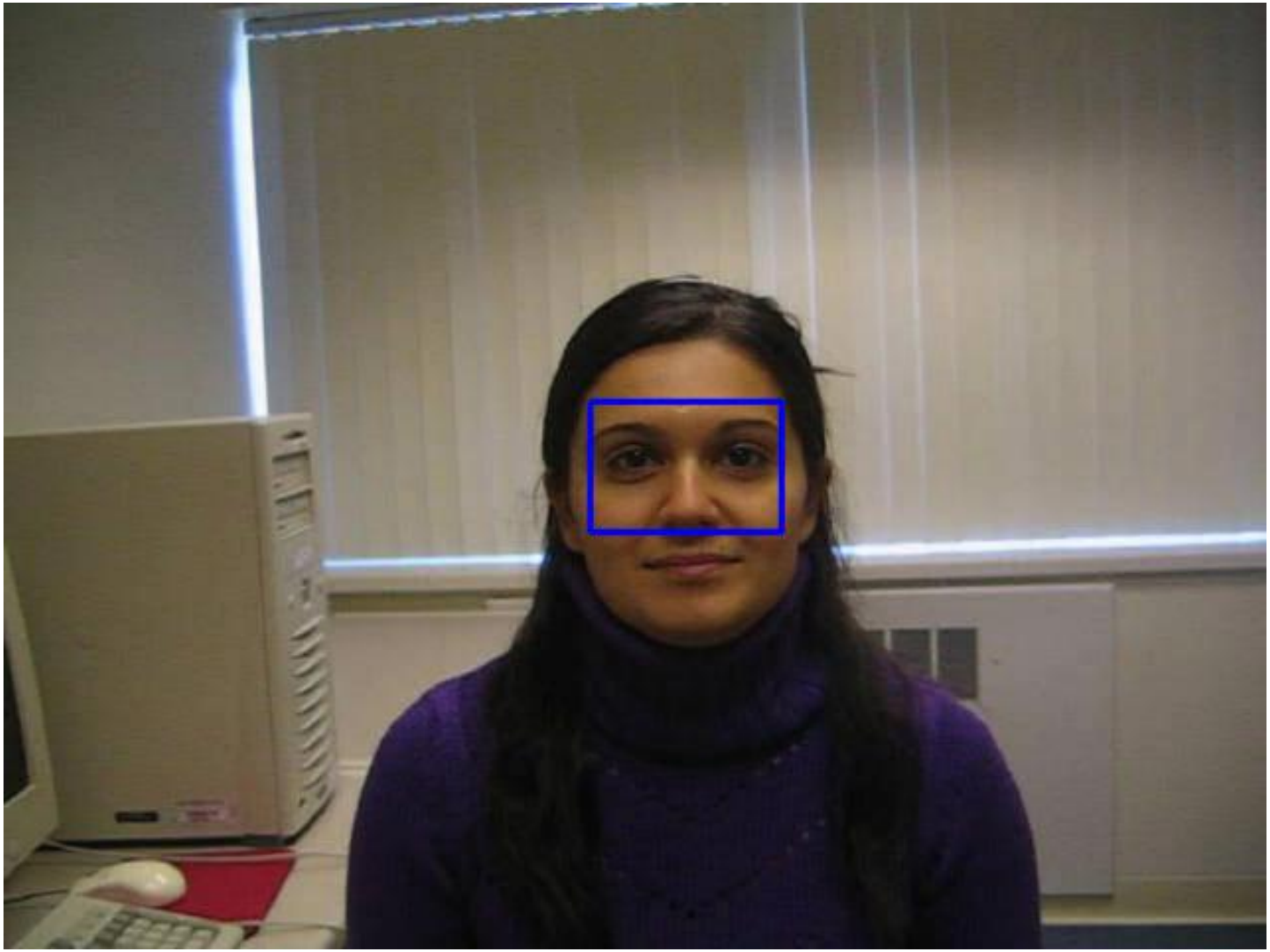




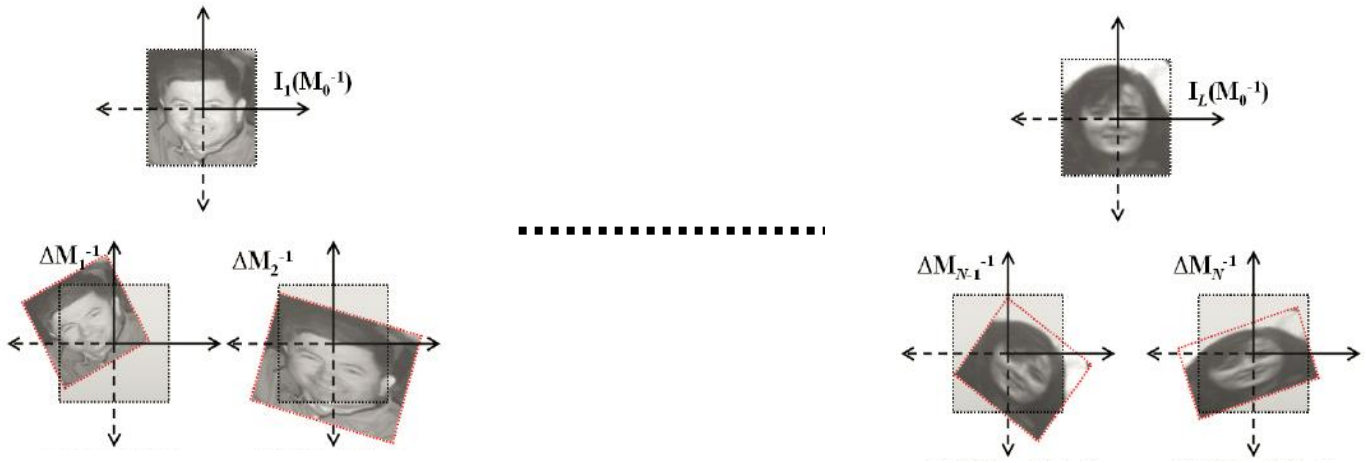
Regression on Manifold



Regression on Manifold



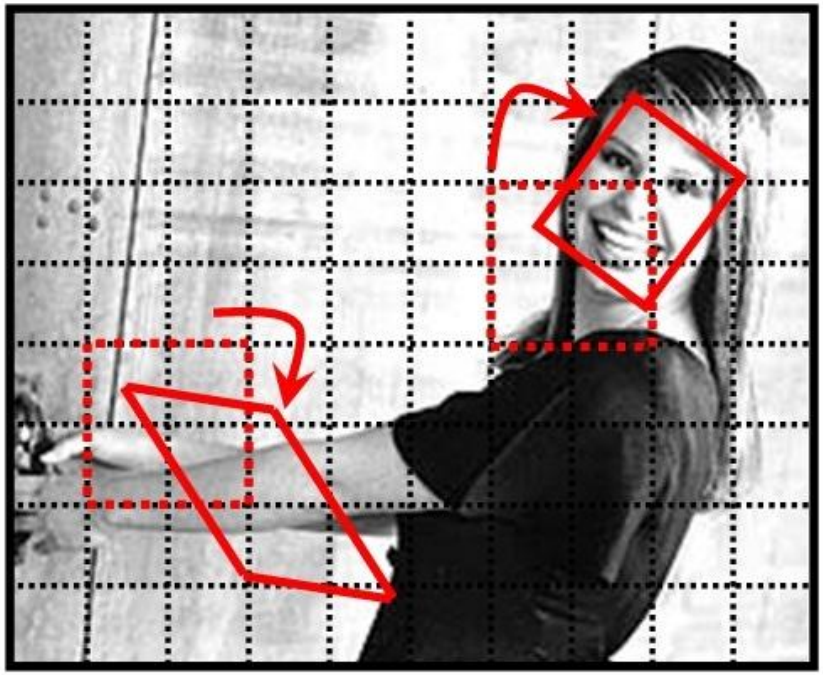
Learning Class Specific Regression



- Tracking function is learned on a training set of n random affine transformations applied to an object class
  - A class specific tracking function (e.g. tracker for faces) is integrated to an existing pose dependent detector

- Training is an offline process and a more complex model can be learned compared to tracking application

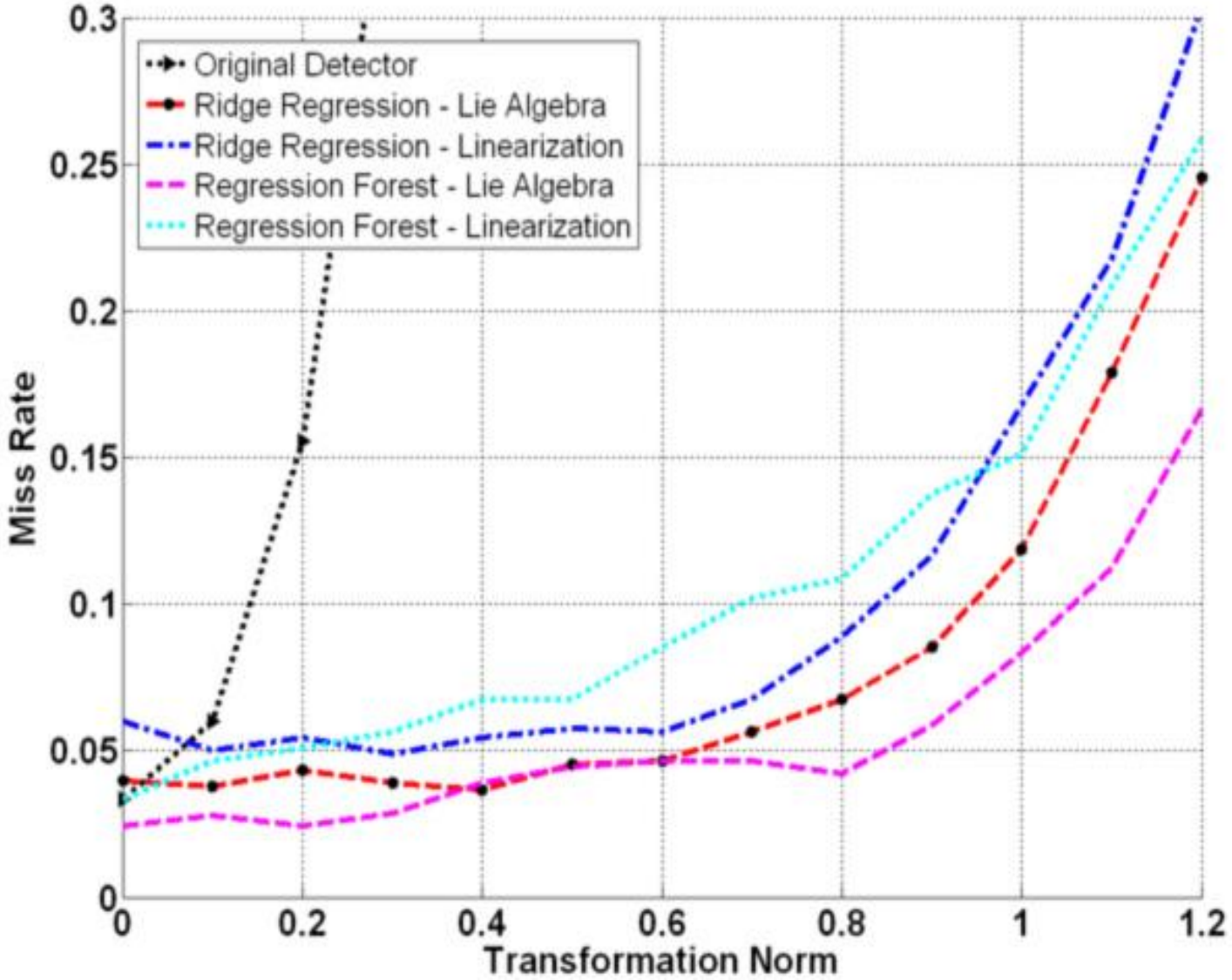
Learning Class Specific Regression



- With a sparse scan of the image, regression finds all locations which resemble the object model
  - The object detector is then evaluated only at these locations

Learning Class Specific Regression

Learning Class Specific Regression



Learning Class Specific Regression

(First row) Regression, (Second row) VJ detector with a dense scan

Last Word...

- Several parameter spaces commonly occur in computer vision problems have Riemannian manifold structure

- Manifold based methods provide major improvements over the existing techniques
  - Graph Laplacian next?

Thanks!

Summary

Not shown

- Segmentation
- Compressive sensing, sparse solvers, dictionary learning
- Fast SVM's, distance learning, kernel PCA, online/incremental learning, scene adaptation
- Lots of tracking methods
- Video classification
- GPU optimized vision
- Bilateral filtering
- Visualization of periodic phenomena

Outline

Part 1: Professional background

Part 2: Snapshots of some past projects (breath)

Part 3: Manifold methods (depth)

Part 1:
Professional Background

Professional Background

- Senior Principal Research Scientist & Project Manager, MERL, MA
  - Hughes Research Laboratories, CA, 1999
  - AT&T Research Laboratories, NJ, 1998

- PhD: Polytechnic Institute of New York University, NY, 2002
  - Segmentation

- Authored 80+ publications and invented 50+ patents
  - Best Paper Runner Up Award CVPR 2007
  - 2 Best Paper Nominations at ICME 2008 & AVSS 2009
- Mentored 30+ PhD students

- Recipient of the R&D 100 Award 2006
- Presidents Award MELCO Japan 2007, Superior Invention Award MELCO Japan 2008, Research Excellence Award MELCO-PUS Japan 2009, Directors Award MERL 2008

Professional Background

- Associate Editor, Journal of MVA, 2006 to present
- Associate Editor, Journal of RTIP, 2004 to present
- Guest Editor, JMVA, SI on Car Navigation & Vehicle Systems, 2010
- EURASIP JIVP, SI on Tracking in Complex Scenes, 2008

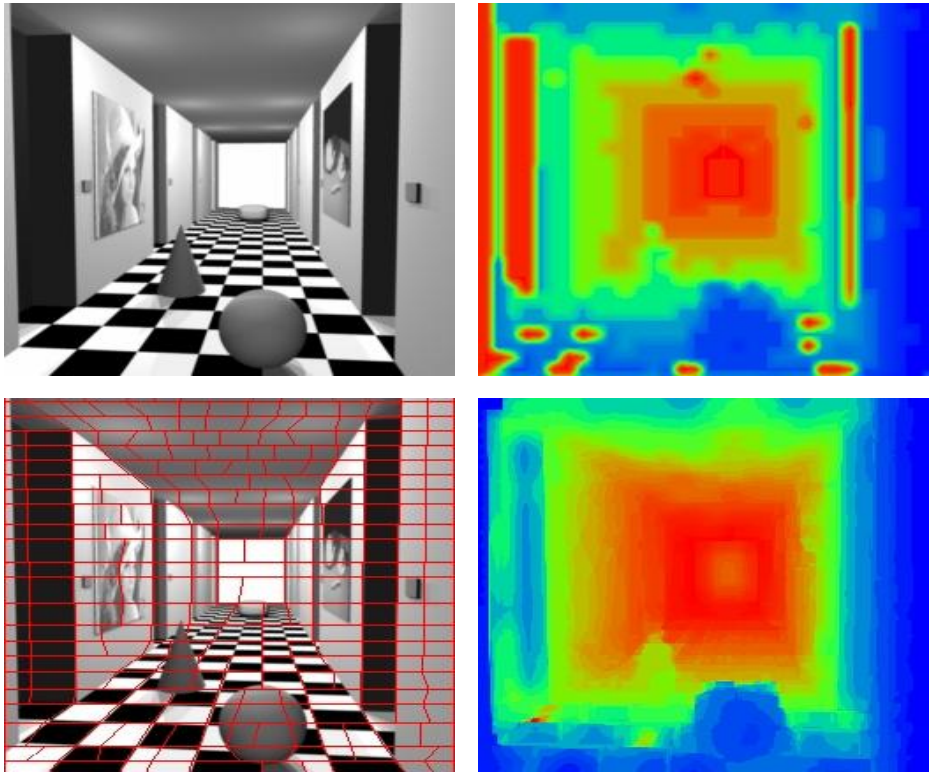
- General Chair, IEEE AVSS 2010
- Advisory Board, IAPR ICPR 2010
- Program Chair, VCIP 2004 to 2008, SPIE RTI 2003 to present
- Track Chair, IEEE ICME, 2007
- Area Chair, IEEE CVPR 2009 and ICPR 2010
- Special Tracks Chair, ISVC, 2009, 2010
- Industrial Liaison IEEE ICCV 2011, Liaison IEEE IV 2009
- Organized 20+ workshops, etc. (OLCV), TPC at 40+ events

- NSF Panels, Robust Intelligence, 2008, 2010

Part 2:
Snapshot of “Some” Projects

Patch Match

- Depth estimation is crucial for robot vision
- Pixel-wise approaches: aperture problems, imposing smoothness on objective function, handling depth discontinuities
- Elegant combination of piecewise planar surfaces on epipolars
- Achieves accurate yet smoother depth maps



Conventional

Patch match

Blind Video Segmentation

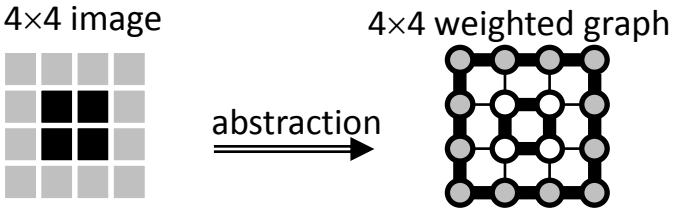
- Automatic segmentation is essential for object based coding
- How to incorporate motion without explicitly computing motion, enforce boundary consistency, and reduce computational load?
- Use 3D spatiotemporal tubes and agglomerative clustering
- Achieves 100x faster segmentation, hierarchical object trees

Motion saliency →

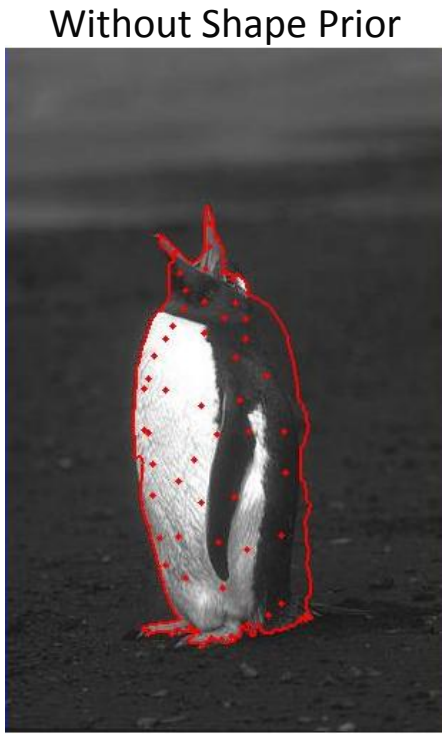
Semi-supervised Segmentation

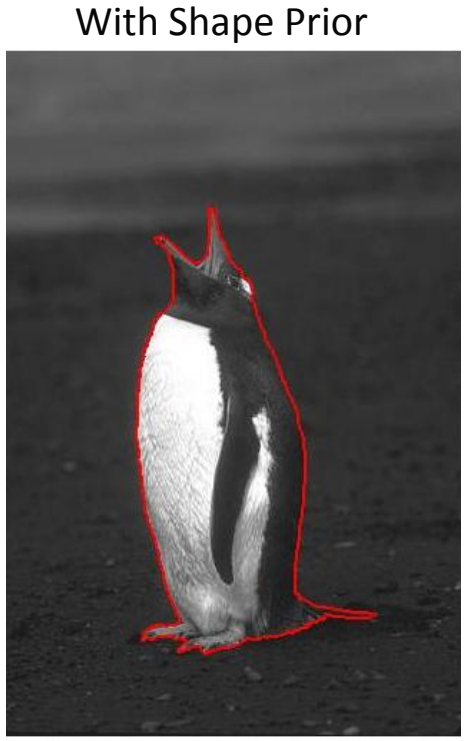
- Generate weights based on image intensities & shape priors
- Build graph Laplacian, solve system of equations for each label
- Assign pixel to label for which it has the highest probability
- Very accurate semi-automatic segmentation tool

Image to a weighted graph



Compute probability that emitted graph front from seeds arrive at pixels





Intrinsic Images

- Application: lighting compensation for recognition, tracking tasks
- Lighting conditions and surface structures are unknown
- Intrinsic images as midlevel decomposition, ML on sparseness
- Achieves moving object detection under drastic lighting changes

Original image

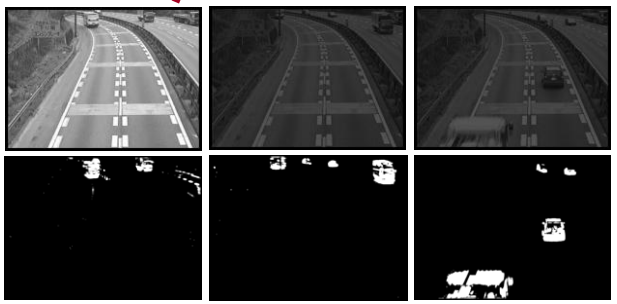
Reflectance map

Illumination map

$I = R \cdot L$

Before lighting change

After lighting change



Intrinsic

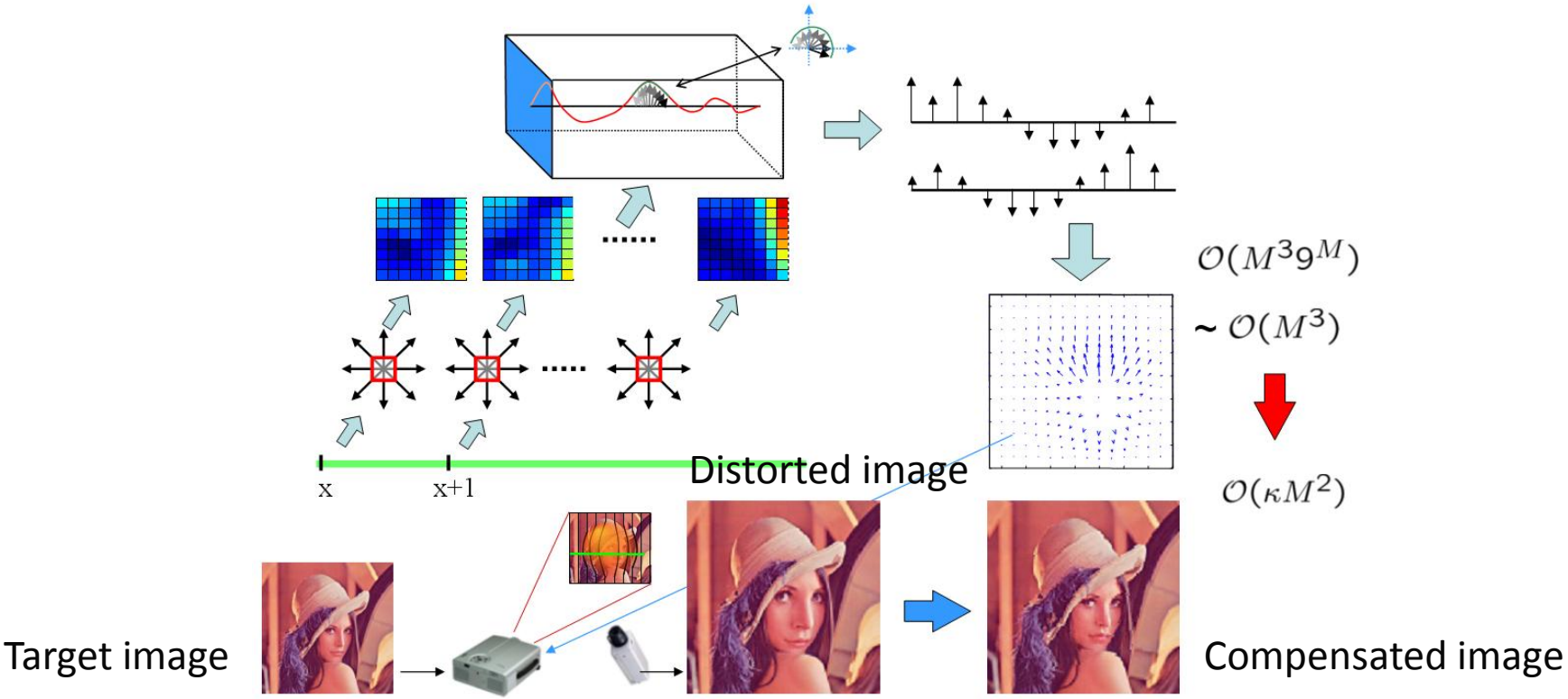
Conventional GMM

Bilateral Filtering

- Bilateral filters smooth redundant texture yet preserve edges
- Joint range & domain filtering, computationally very intensive!
- Integral histogram, power images, Taylor expansion
- O(1) load, fastest bilateral filter in 2008 (200 fps @ 1MB GPU)

Camera Distortion Correction

- Application: projection of video onto dynamic surfaces
- Real-time calibration is required, conventional solutions are slow
- Warping estimation by dynamic programming in motion space
- Achieves load reduction from “exponential cubic” to “quadratic”

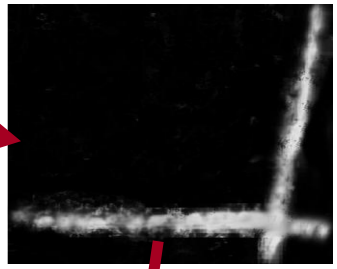


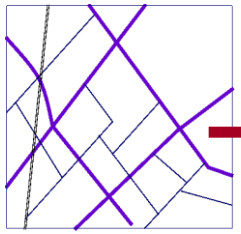
Road Detection

- Geo-mapping, navigation, aerial camera position refinement
- Low-resolution, significant appearance variance, good features?
- Recursive propagation, orientation cues, multi-layer neural net
- Improves alignment accuracy more than 5x , speed 1000x

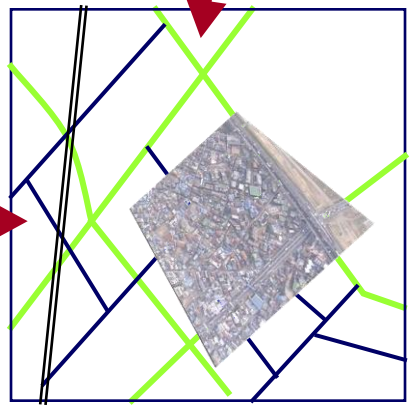
Aerial image

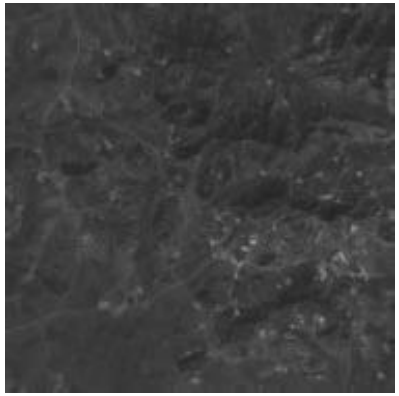
Detected highways



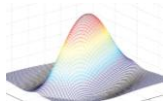


Vector map





Satellite image



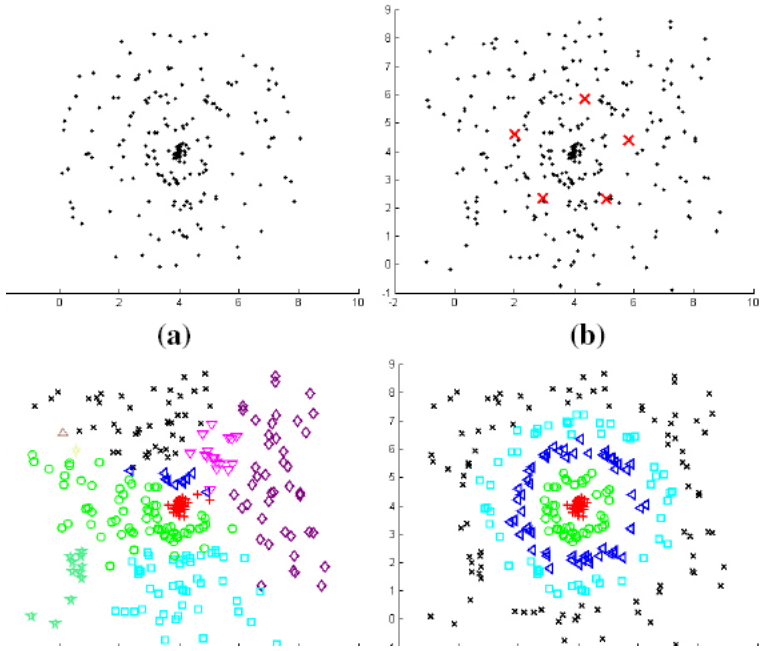
Detected roads

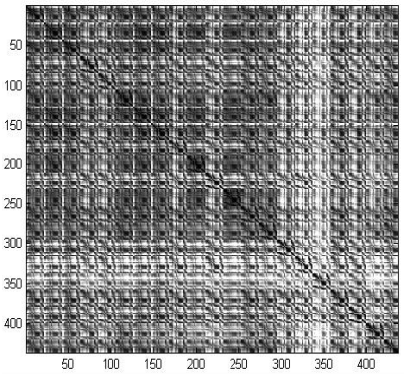
Weakly-Supervised Kernel MS

- Clustering is a core task for many vision problems
- Fully automatic methods would not obey underlying manifold
- Subspace projections on must link constraints in kernel mean-shift
- First density estimator that can incorporate weakly labeled data

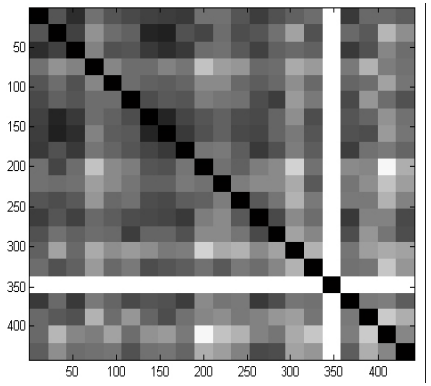
Kernel-MS

WS K-MS





Kernel-MS: 63%



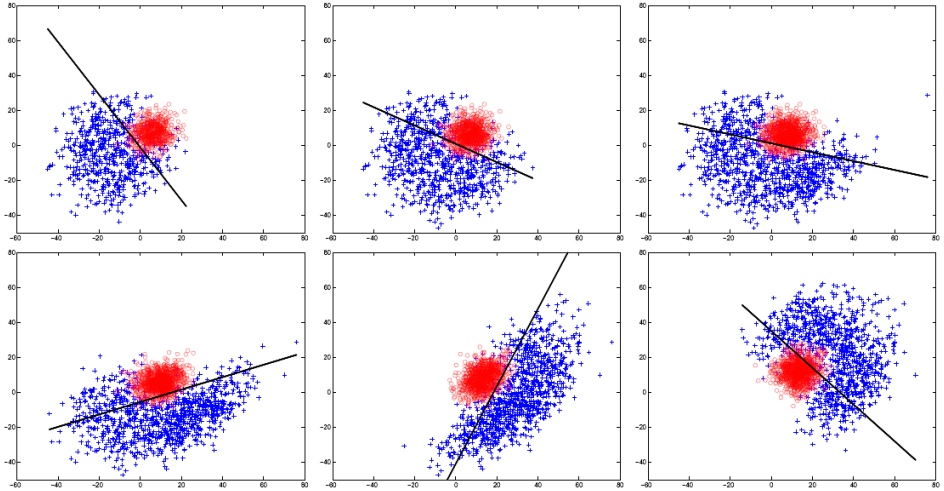
WS K-MS: 100%

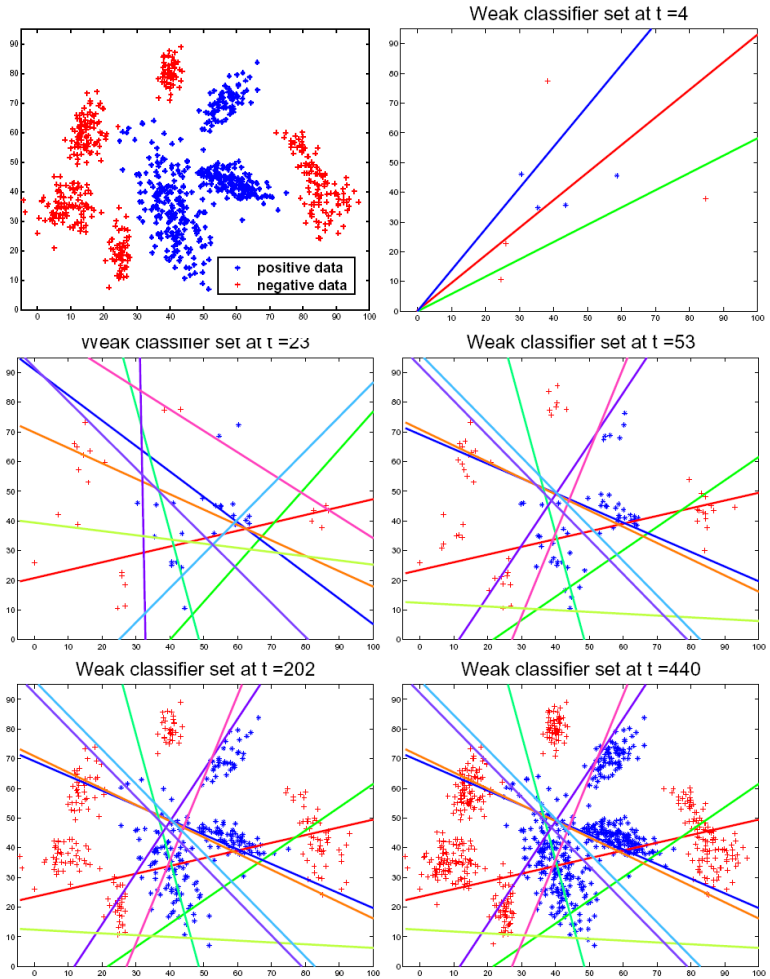
Online Boosting of LC Ensembles

- An efficient method for the adaptation of linear weak classifiers
  - Instead of replacing, weak classifiers themselves are modified
  - Avoids the issue of how many weak classifiers to be replaced or which fast search algorithm to use
- Zero memory

Adaptation of linear classifier:

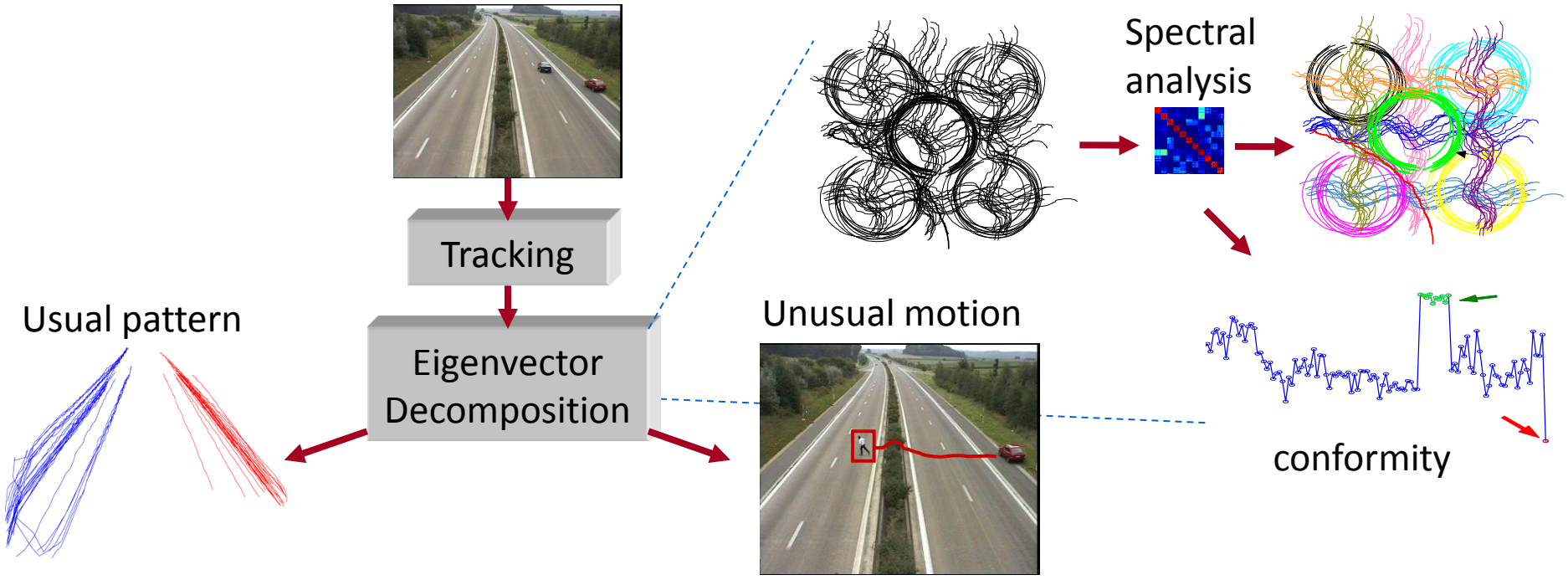
Top: without forgetting, Bottom : with forgetting





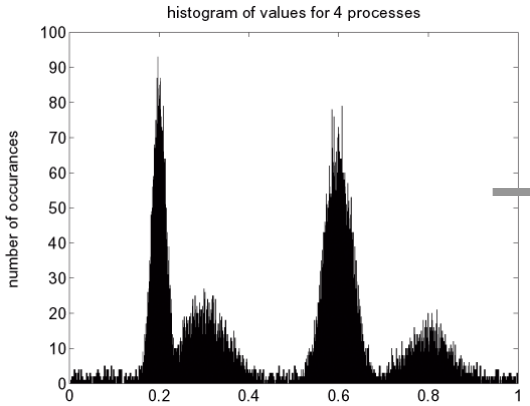
Event Detection

- Analytics is the ultimate goal of surveillance systems
- Without requiring higher level semantics and user interaction, how to distinguish usual from unusual?
- Cross-fitness trajectory alignment, spectral clustering, conformity
- Enables automatic detection of usual & unusual motifs

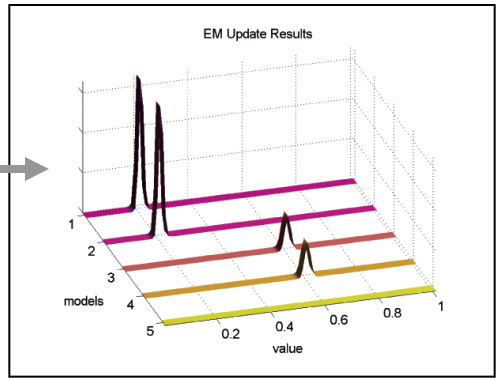


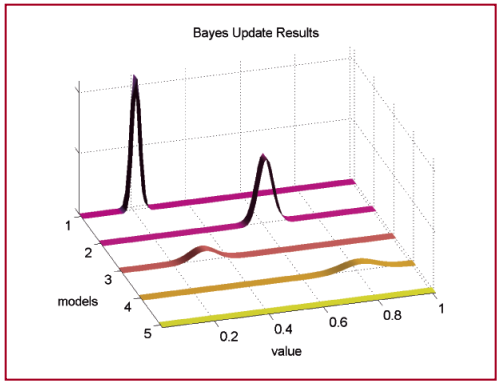
Online Bayesian Update

- Background learning is essential for static camera surveillance
- GMM with EM update fails for complex scenes blending models
- Instead of mixture, use layers of multivariate Gaussians
- Bayesian update preserves multimodality, automatically determines necessary layers



Sampled distribution

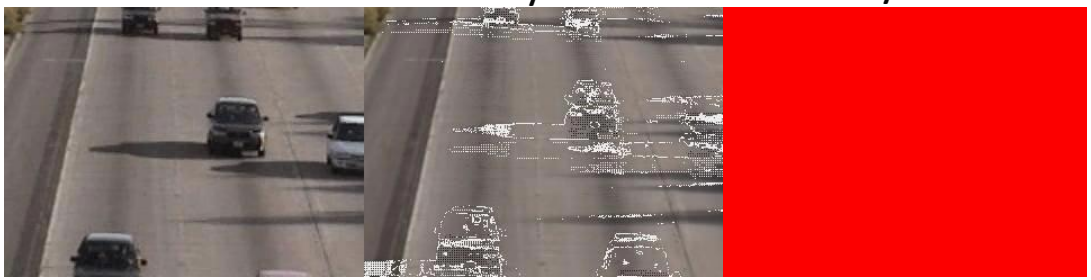




Layer 1

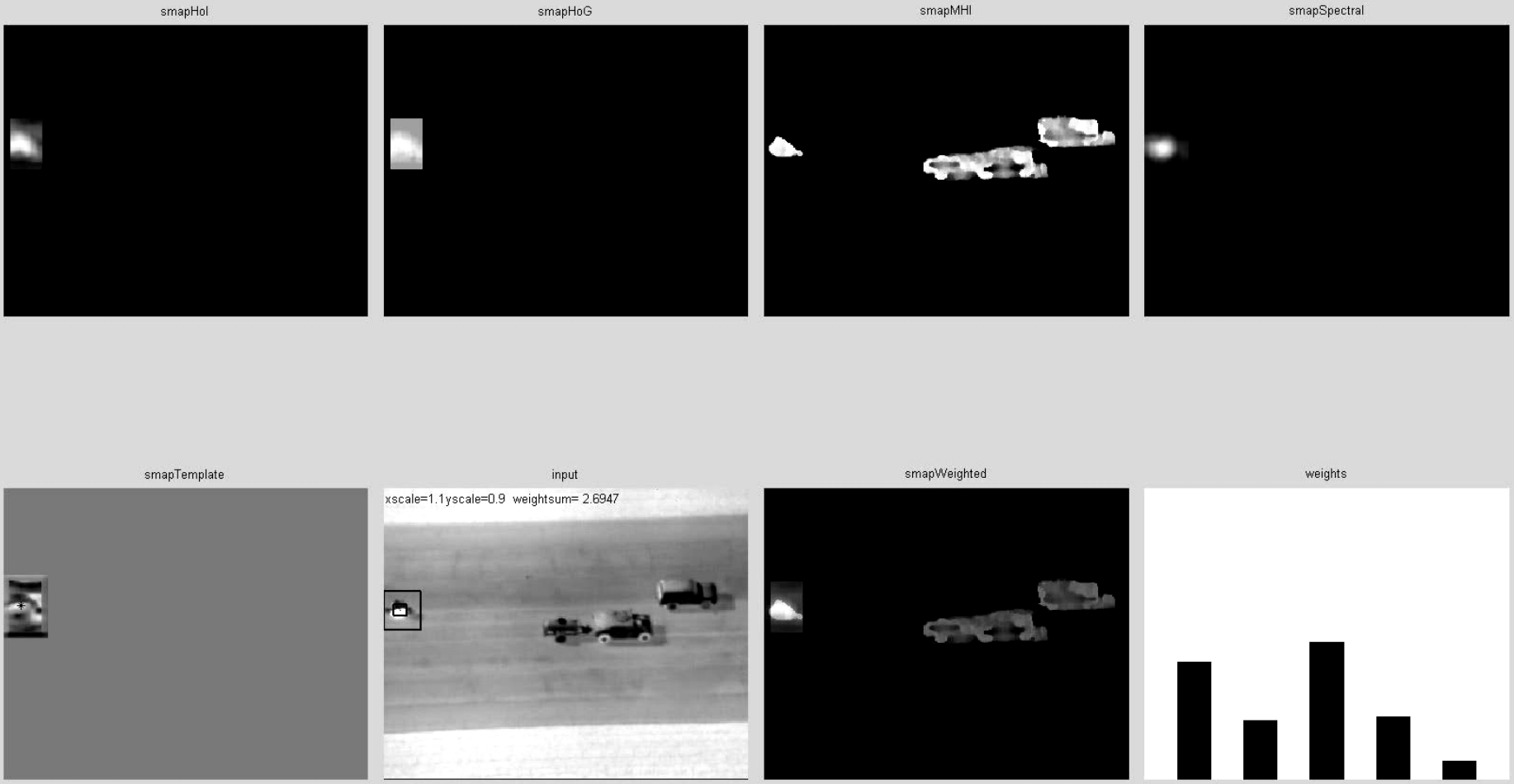
Layer 2

Traffic video



Model Update on Manifold

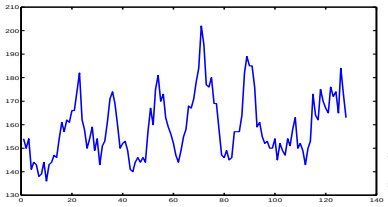
Intensity histogram | Orientation | Motion history | Saliency



Template similarity | **Final result** | Likelihood map | Fusion weights

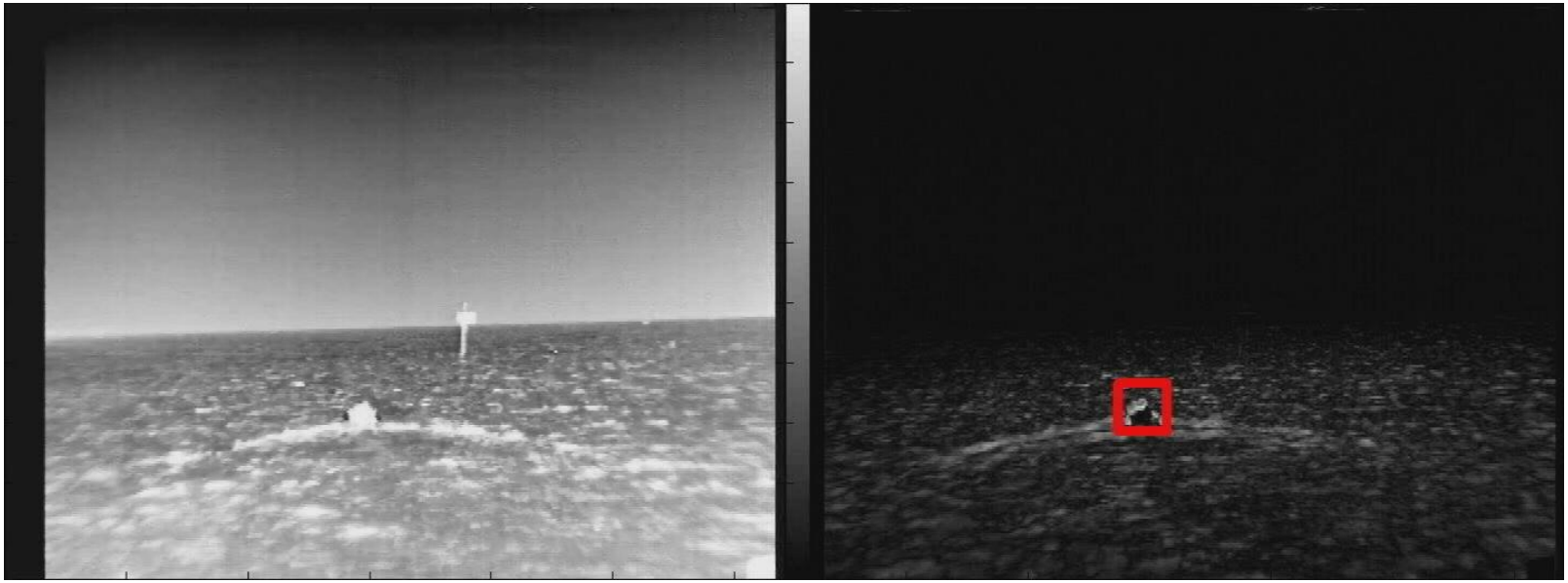
Cyclo-Stationary Scene Learning

- Certain backgrounds are in motion: sea, trees, mirage, turbulence
- Intensity models fails for such scenes causing false detection
- Mixture modeling of frequency instead of magnitude
- Accurate detection of moving objects, 4x less errors



$$x[t] = \sum_{k=0}^{N-1} a[k] e^{i2\pi kt/N}$$

Thermal IR, seaside with a boat



Left Behind Items

- Finding left-behind/removed objects is a main surveillance task
- Tracking may not work always, complementary method is needed
- Dual background are adapted at different learning rates
- Real-time detection with customizable settings, robust method adapts to scene changes

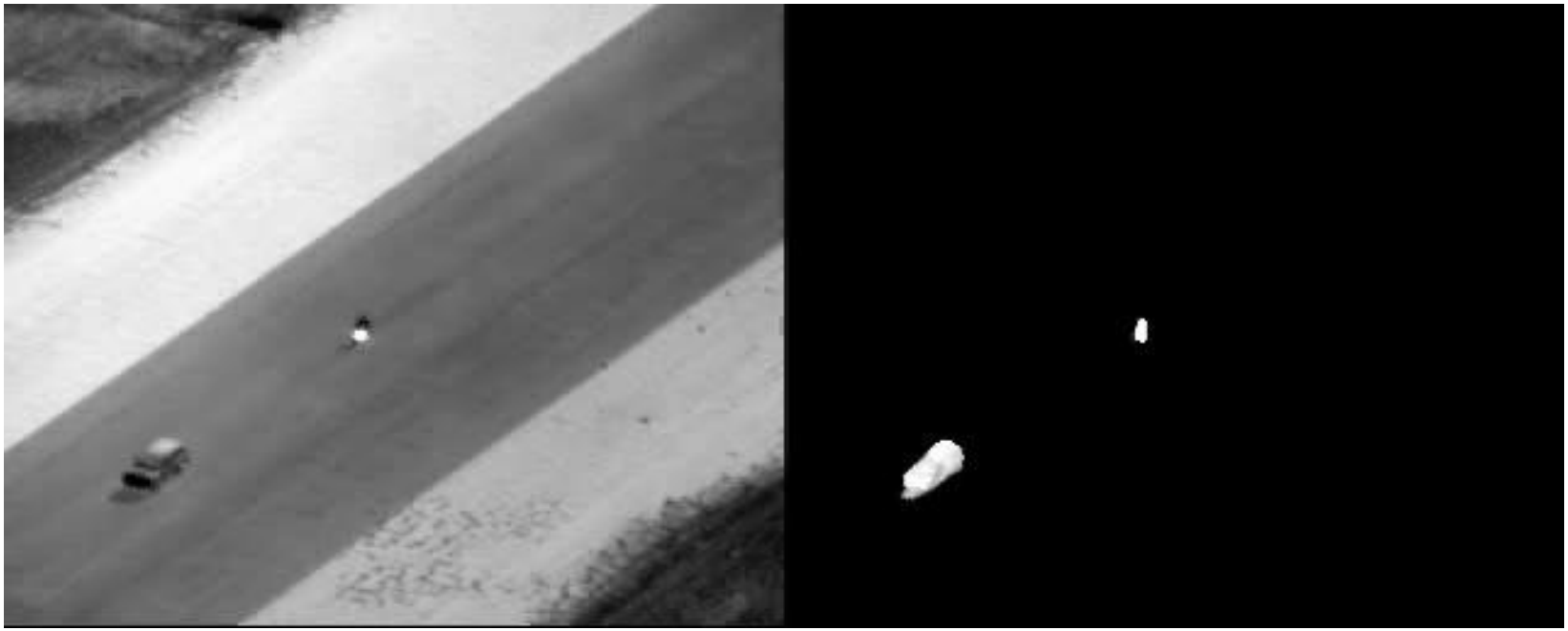
Multi-kernel mean shift

Dual background

Detection for Moving Object/Camera

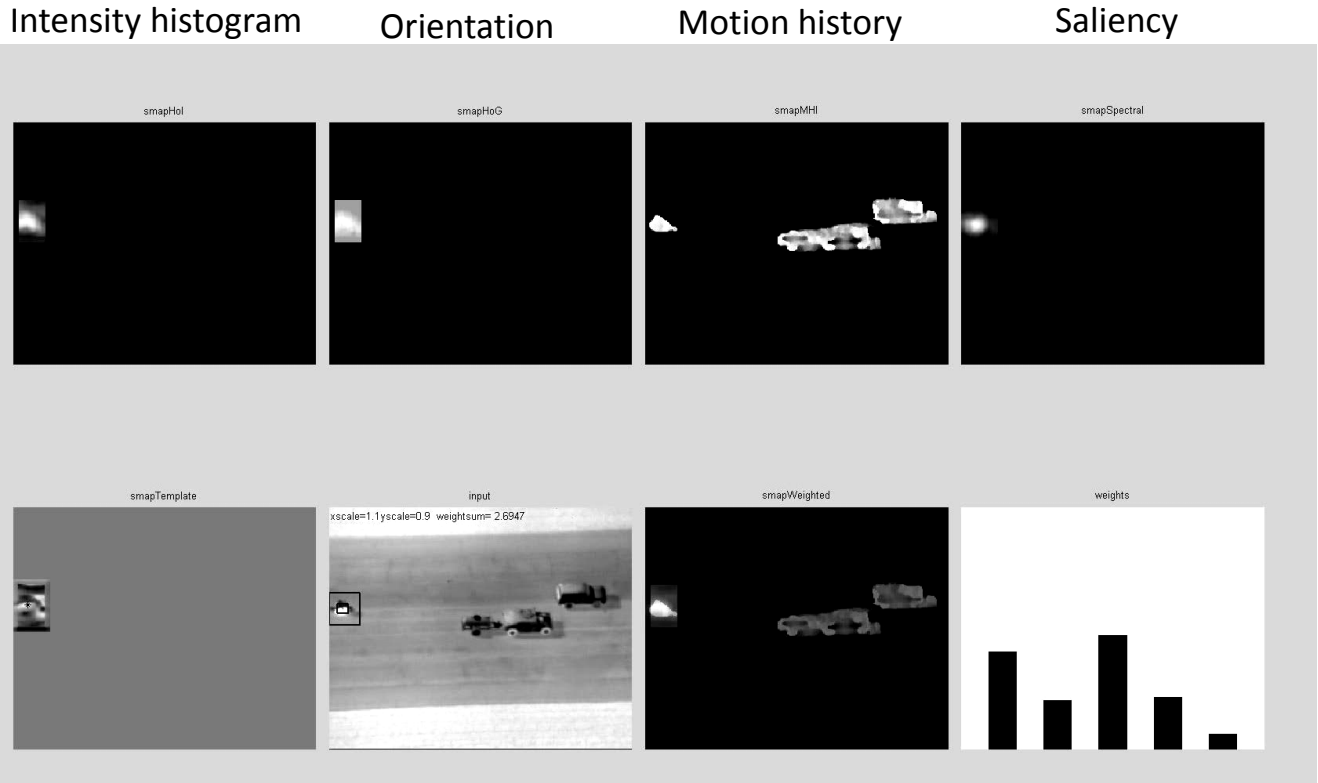
- Application: UAV's and other aerial platforms
- Updating a scene model for moving cameras is costly
- Forward/backward motion history with RANSAC motion estimation
- Computationally efficient, robust, works for low-quality data

Detection result



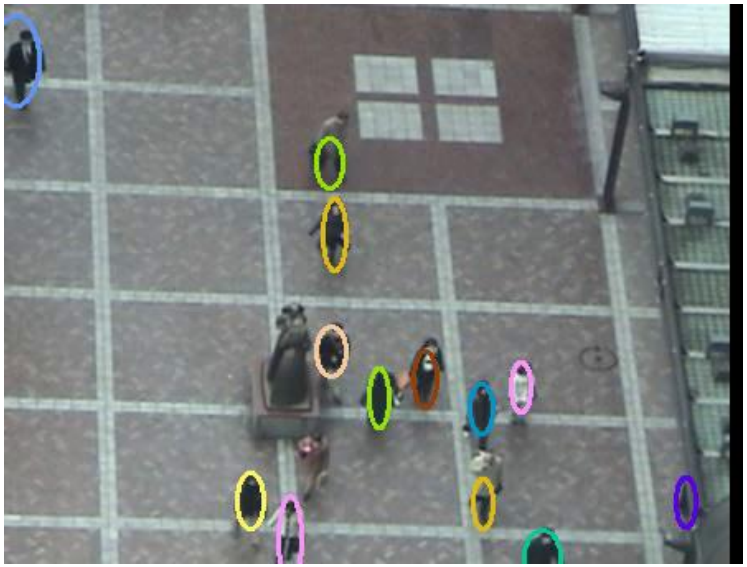
Tracking for Moving Object/Camera

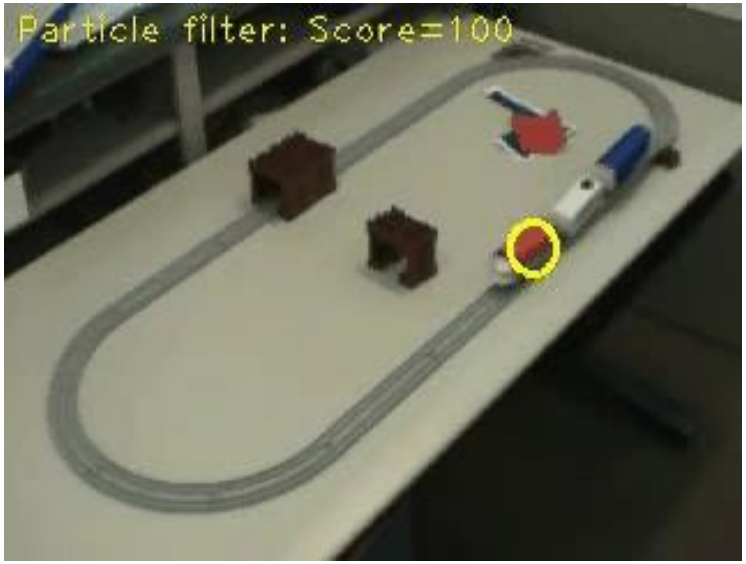
- Application: UAV's and other aerial platforms
- Tracking in low-quality video, small target size is challenging
- Density estimator on classifier scored likelihood surface
- Automatically selects the most discriminative features

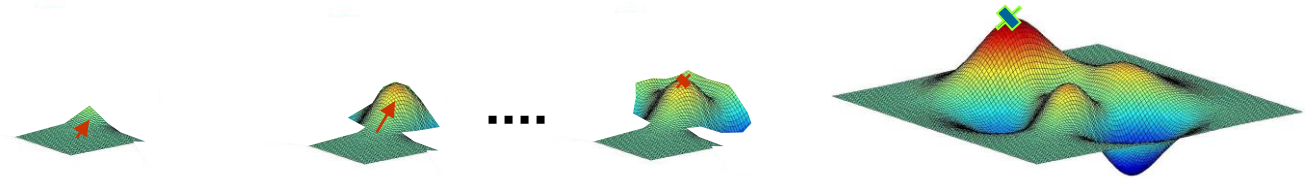


Tracking for Static Cameras

- Core technology for surveillance systems
- Occlusions, appearance variation, shadows, lighting changes
- Multi-kernel mean-shift, particle filter, ensemble tracker
- Fast, robust trackers (5msec/object), low frame rate tracking

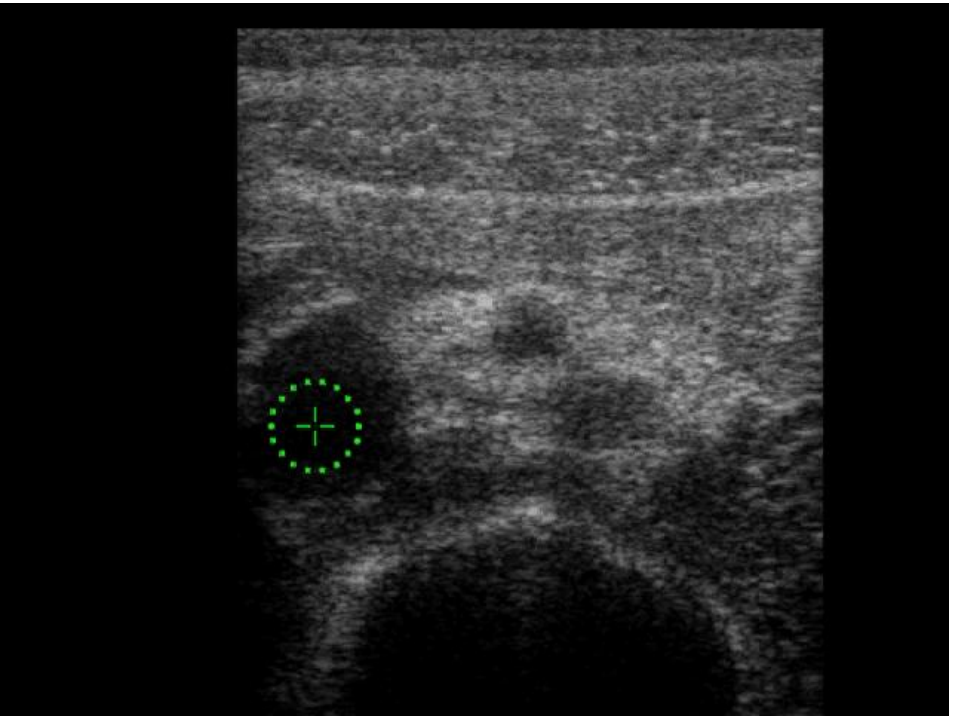






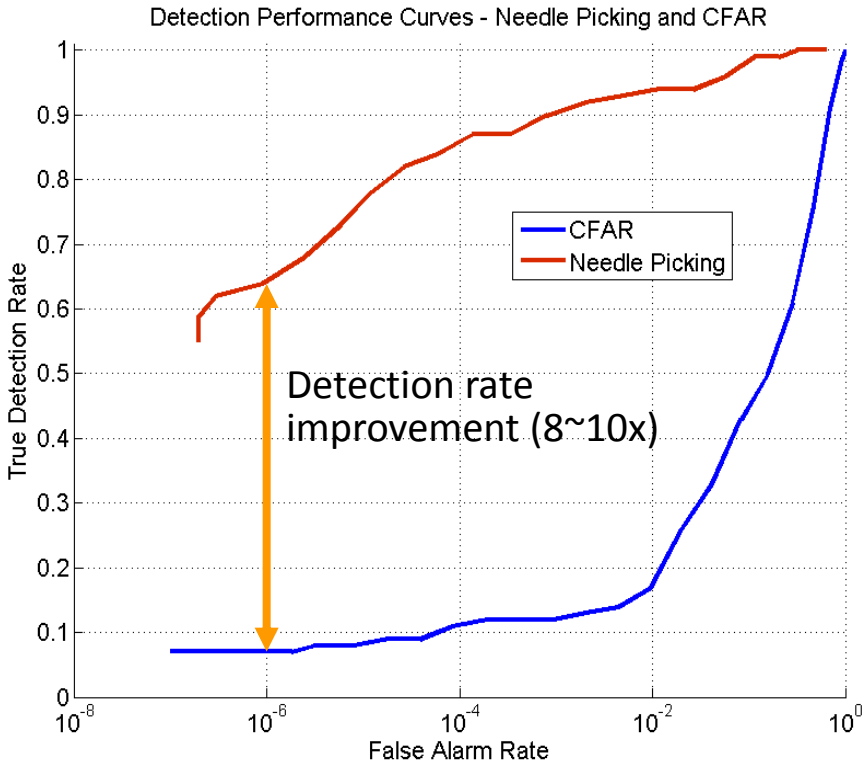
Tumor Tracking

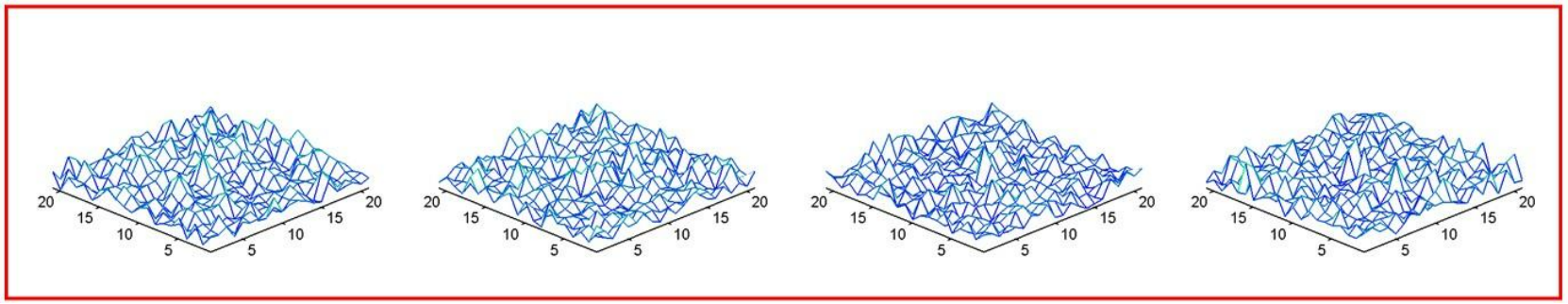
- Particle beam therapy is the most advanced tumor treatment
- Tumors dislocate due to breathing, etc. during treatment
- Non-intrusive approach, regressed & graph propagation methods
- Very accurate (<2mm error), computationally very fast (<30 msec) methods for visible & invisible tumors in ultrasound & X-ray videos



Radar Target Detection

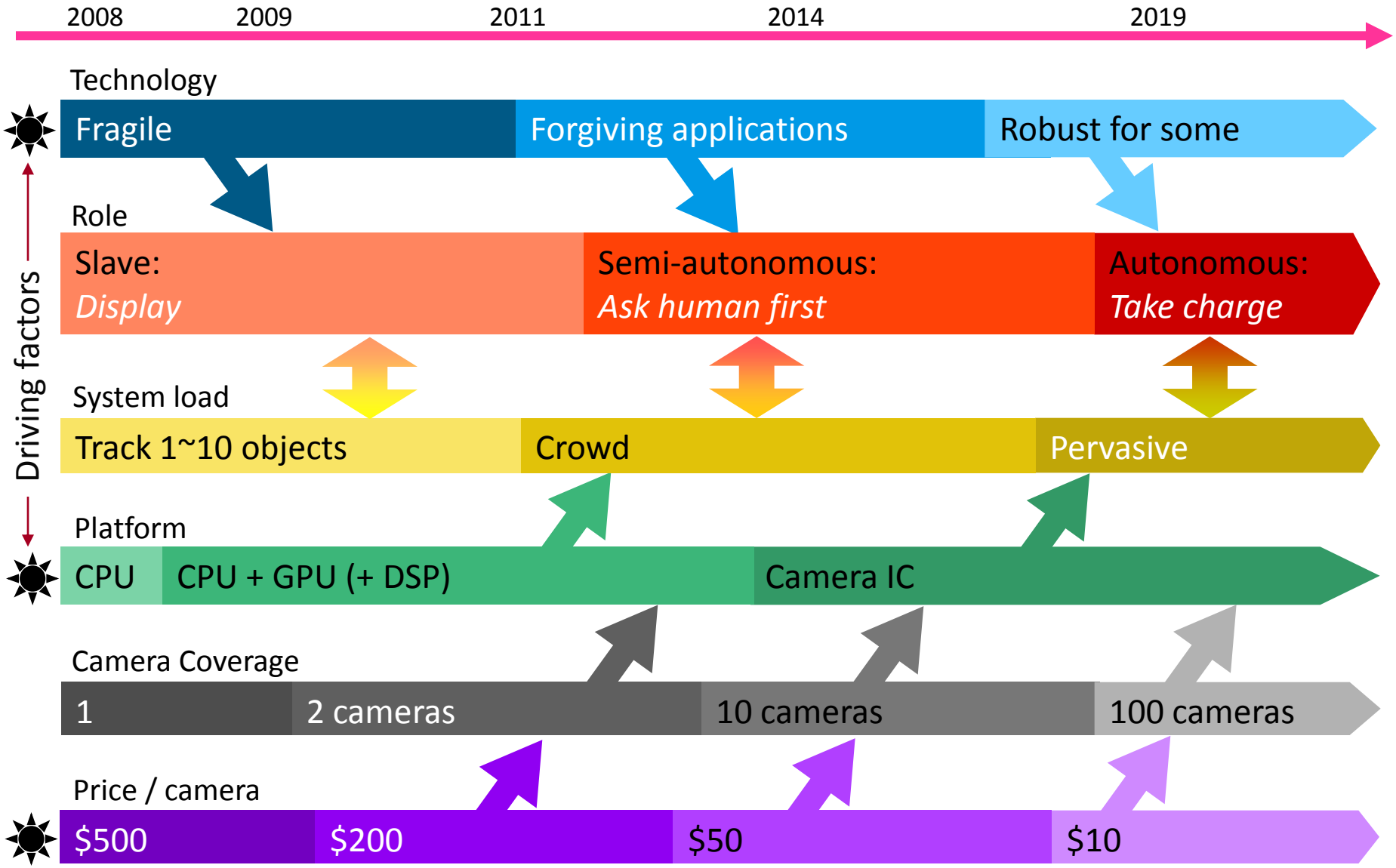
- A sampling based hypothesis verification approach
- Any motion can be incorporated, multiple targets can be detected
- No assumptions on noise & clutter distributions
- Very fast, complexity is easily scalable wrt performance
  - Much faster than Bayesian, Viterbi. Faster than Particle filter, P-MHT
- 8x improvement
  - 10^{-6} false alarm rate for SNR < 7dB





Supplementary Slides

Future Trends

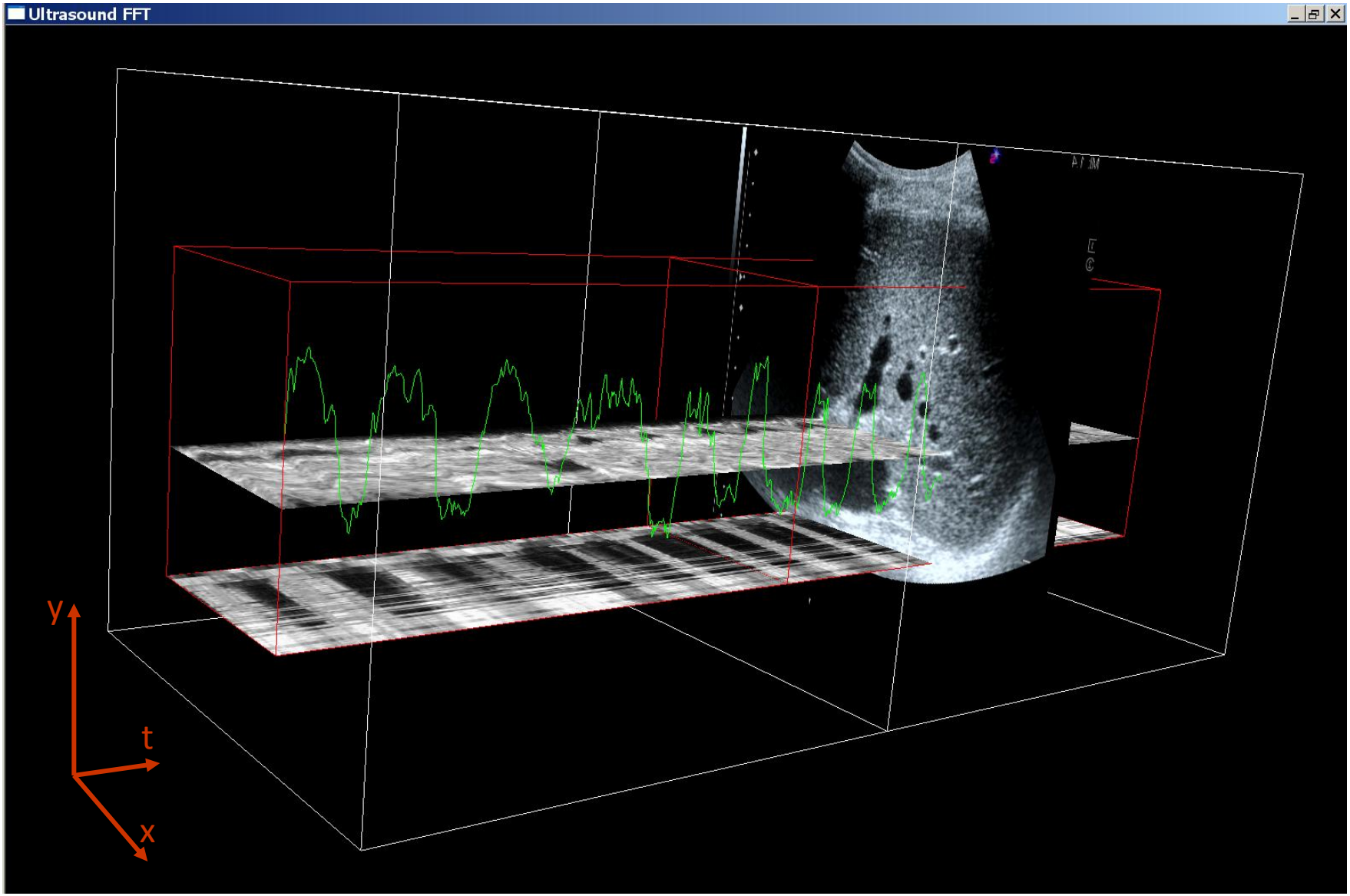


Tracking Comparison

- Methods differ in terms of:
  - Speed, regularity, type of the motions that they can track
  - Single, joint, crowd modes
  - Partial, full occlusion handling
  - Model update mechanisms
  - Computational load

	Optical Flow	Kalman-MHT	Mean-Shift	Particle filter	Regression
advantages	•Low complexity •Needs only pixel-wise priors	•Can overcome occlusions • Compensate for erratic motion	•Non parametric •Low complexity	•Enables tracking in multiple states •Handles occlusion and fast objects	•Can track the pose of an object •Very low complexity
disadvantages	• Very sensitive to clutter and occlusions • Can only detect small motion	• Limited to linear motion • Makes strong Gaussian assumptions	•Sensitive to occlusions •Requires object window to overlap between consecutive frames	•Computationally very expensive •Robustness highly depends on the likelihood function	•Motion model should be parametric

Visualization of Periodic Phenomena



Model Update on Manifold

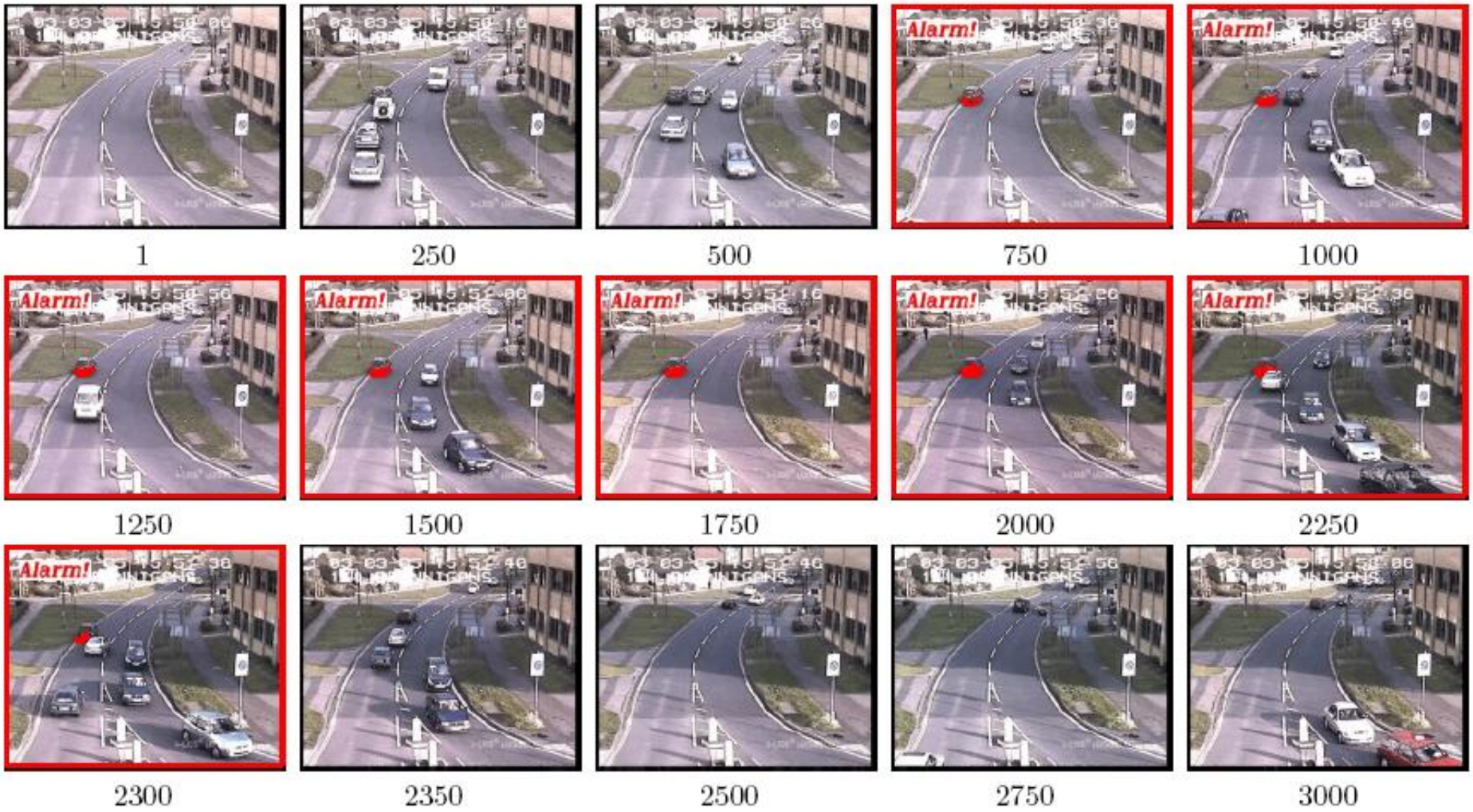
Mean-Shift

Covariance Tracker

- Histograms are difficult to populate for small objects

Model Update on Manifold

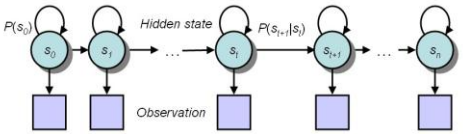
Left Behind Items



Congestion Estimation

- Main application is traffic surveillance and control
- 100's of videos should be processed in real-time
- Use compressed video, DCT/MV features, HMM ML estimator
- Achieved 97% accuracy with 10 seconds latency, 10 msec/video processing speed, geometry independent

Streaming data



Parser

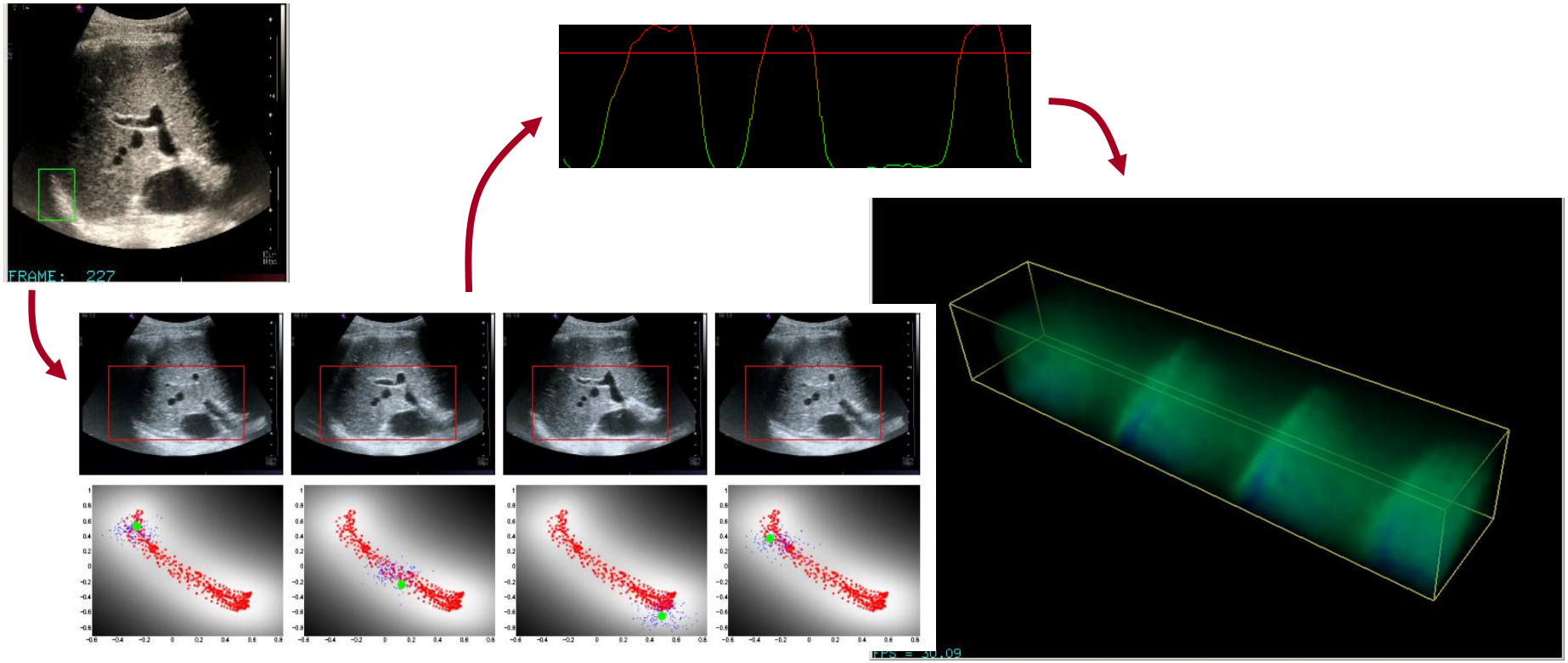
Set of GM-HMMs

ML Classifier

Clogged
Heavy
Mild
Open
Empty

Current Status

Offline training
(setup independent)

Visualization of Periodic Phenomena

- Certain systems require operator to perceive periodicity
- Watching video comprising multiple motions is not comprehensive
- Graph and FFT based representation, Markov state-estimators
- Effective visualization, 10:1 preference on user study