Visual Object Recognition

State-of-the-art techniques and performance evaluation

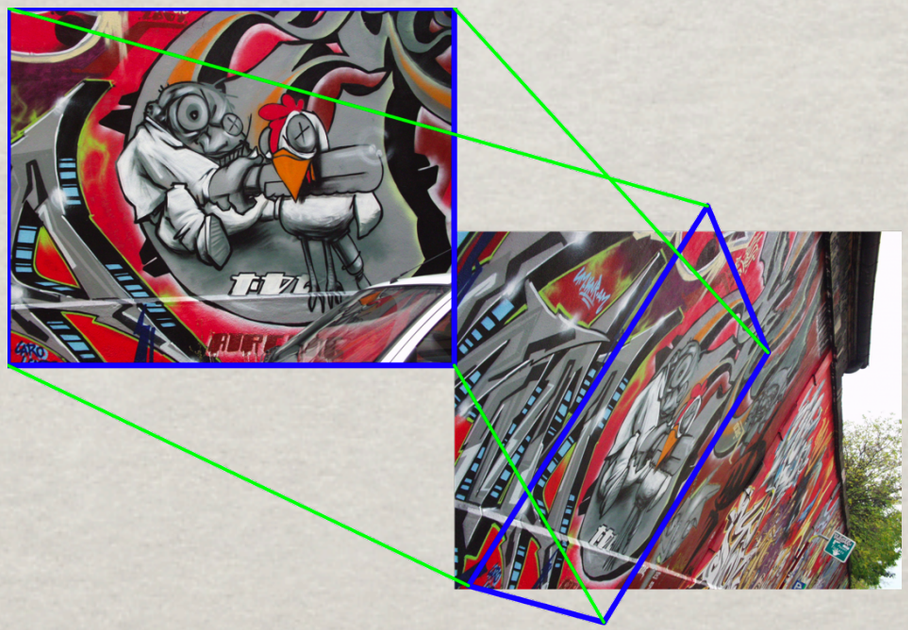
Lecture 7: Performance Evaluation

- Detector: **Repeatability Tests**
- Descriptor matching: **Inlier frequency curve**
- Classifier: **ROC and Precision-Recall curves**
- Discussion of exam and evaluation

Repeatability Tests

- Used for evaluating feature detectors. Article #2, Mikolajczyk IJCV'06.

- Known geometric transformation between two views can be used to check if the same region is detected in two images.



Repeatability Tests

- Example: Homography

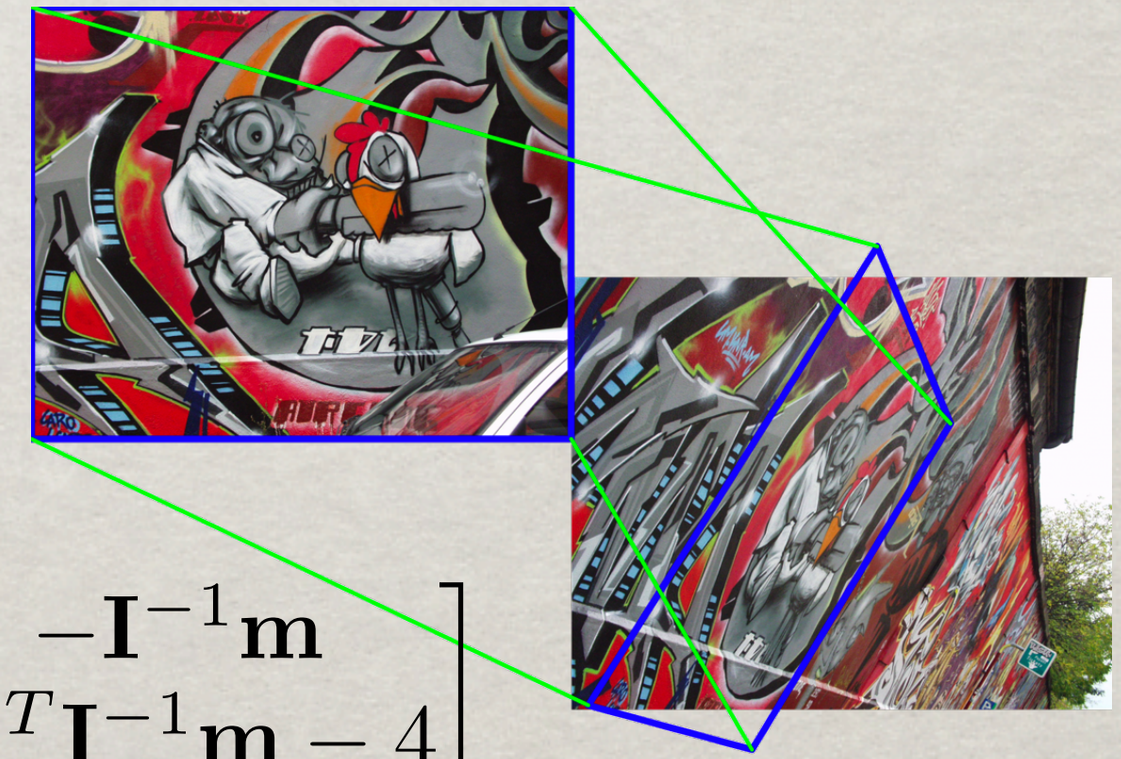
A point **x** should be transformed to a point **x'** according to:

$$\mathbf{x}' = \mathbf{H}\mathbf{x}$$

In reality we detect regions:

$$\mathbf{x}^T\mathbf{C}\mathbf{x} \leq 0$$

$$\mathbf{C} = \frac{1}{4}\begin{bmatrix} \mathbf{I}^{-1} & -\mathbf{I}^{-1}\mathbf{m} \\ -\mathbf{m}^T\mathbf{I}^{-1} & \mathbf{m}^T\mathbf{I}^{-1}\mathbf{m} - 4 \end{bmatrix}$$



Repeatability Tests

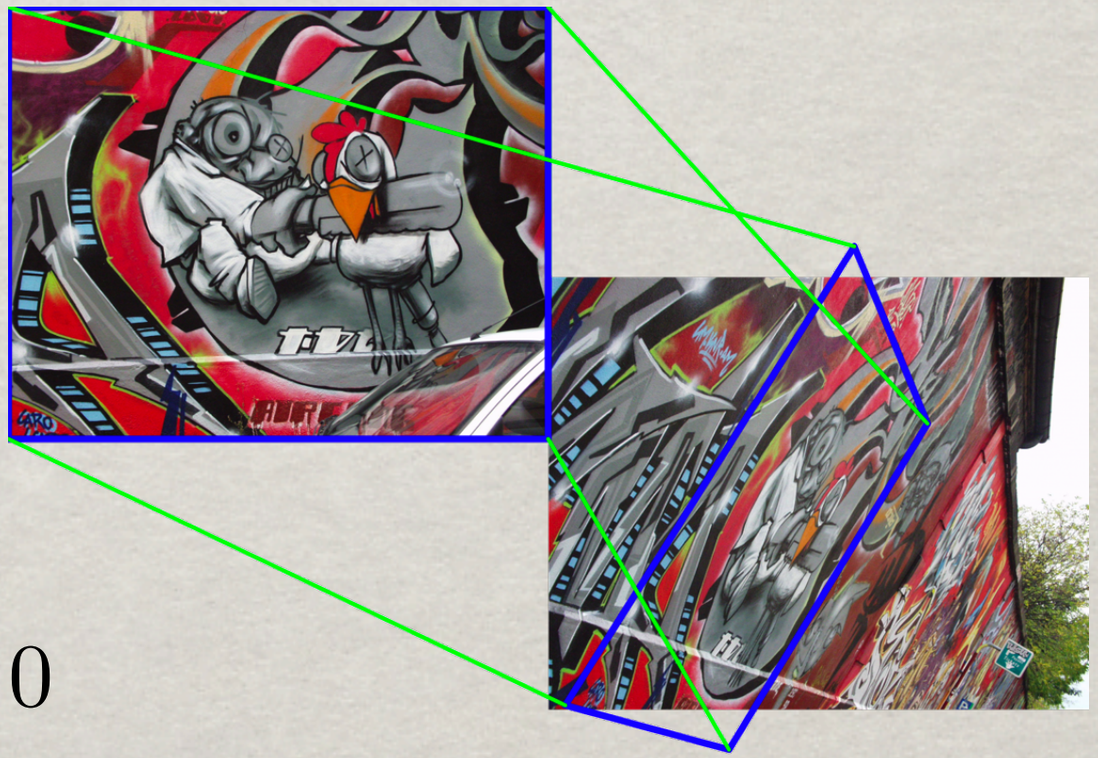
- Example: Homography

An elliptic region **C(m,I)** should be transformed to a region **C'(m',I')** according to:

$$\mathbf{C}' = \mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1}$$

Can be derived from perimeter equation:

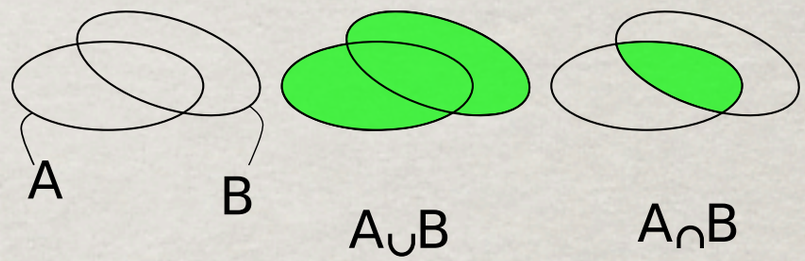
$$\mathbf{x}^T\mathbf{C}\mathbf{x} = 0$$



Repeatability Tests

1. Compute overlap error:

$$\epsilon = 1 - \frac{\text{area}(A \cap B)}{\text{area}(A \cup B)}$$

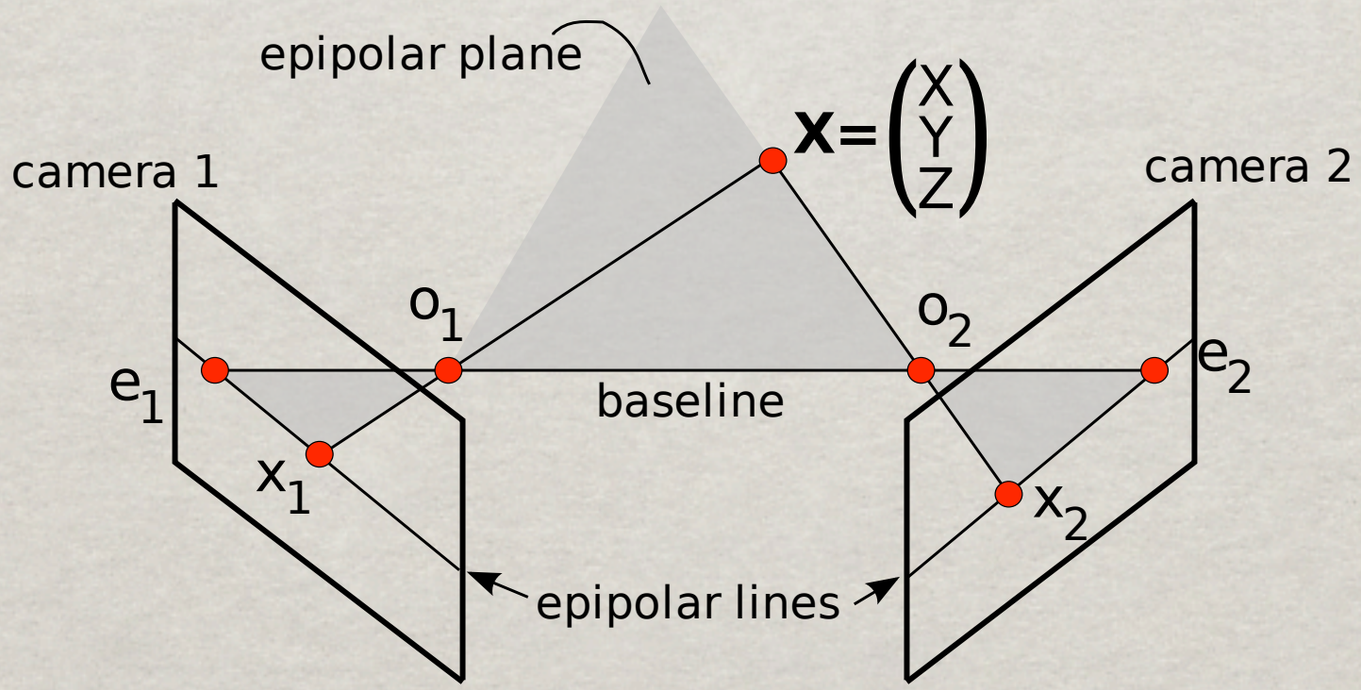


2. Assign 1-to-1 correspondences from image 1 to image 2. (Combinatorial problem if nested regions are detected)

3. repeatability = correspondences (with $\epsilon \leq$ `thr`) divided by #features (in mutually visible region)

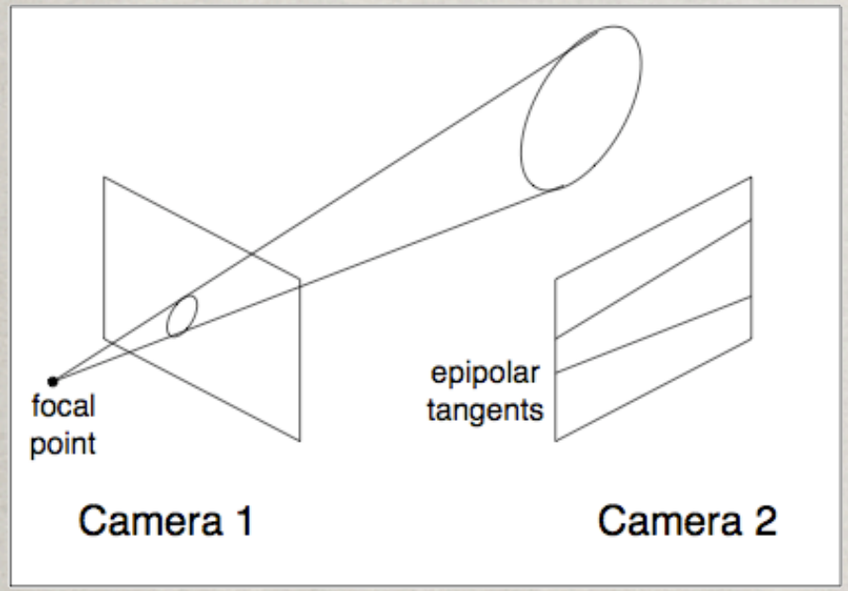
3D Repeatability Tests

- Using generalisation of overlap error to 3D correspondences (Forssén&Lowe ICCV'07)
- Using epipolar geometry, and specifically epipolar tangents.



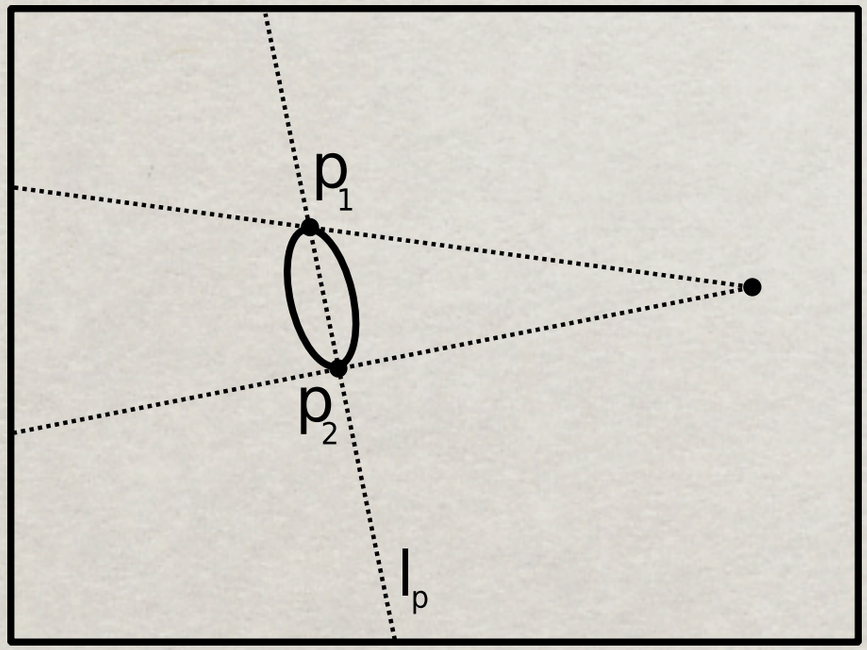
3D Repeatability Tests

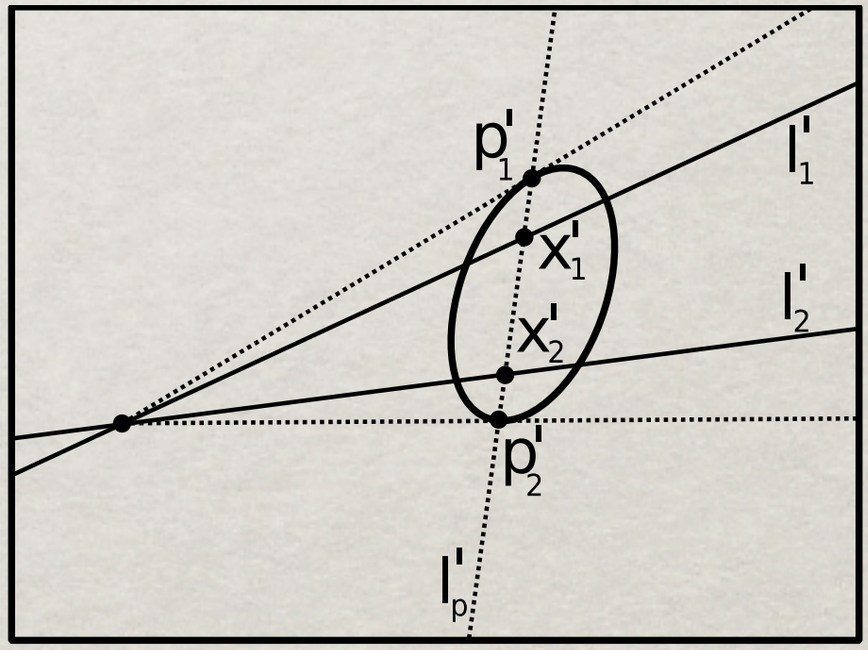
- Epipolar tangents



3D Repeatability Tests

- Measure overlap of tangents and projected epipolar tangents.





$$\epsilon = 1 - \frac{\max(0,\min(x_h,p_h)-\max(x_l,p_l))}{\max(x_h,p_h)-\min(x_l,p_l)}$$

Repeatability Tests

- Repeatability measures probability that a feature will be detected again.
  P(detection|visibility)
- Repeatability is not useful for non-rigid objects/categories.
  As a geometric constraint is used.

Correspondence Count

- A complementary statistic is to simply count the number of correspondences (skip division by number of detected features).

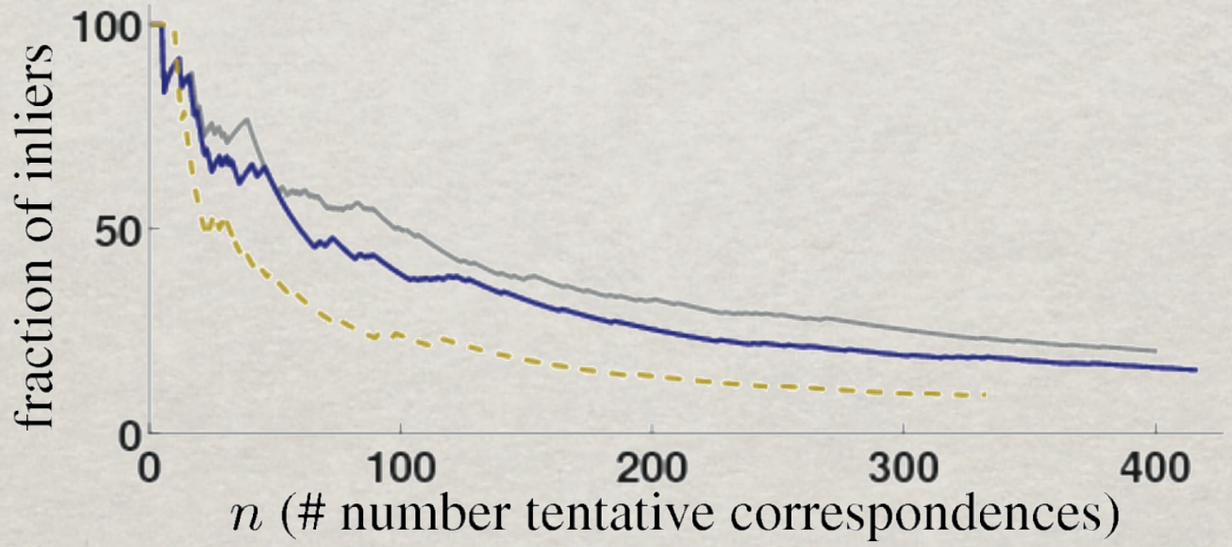
- Good for object recognition:

  If each feature match casts a vote, the probability of a cluster forming by chance is low, so outliers can be tolerated.

- Also: All hypothesis generation(HG)+verification schemes. HG costs only time.

Inlier Frequency Curve

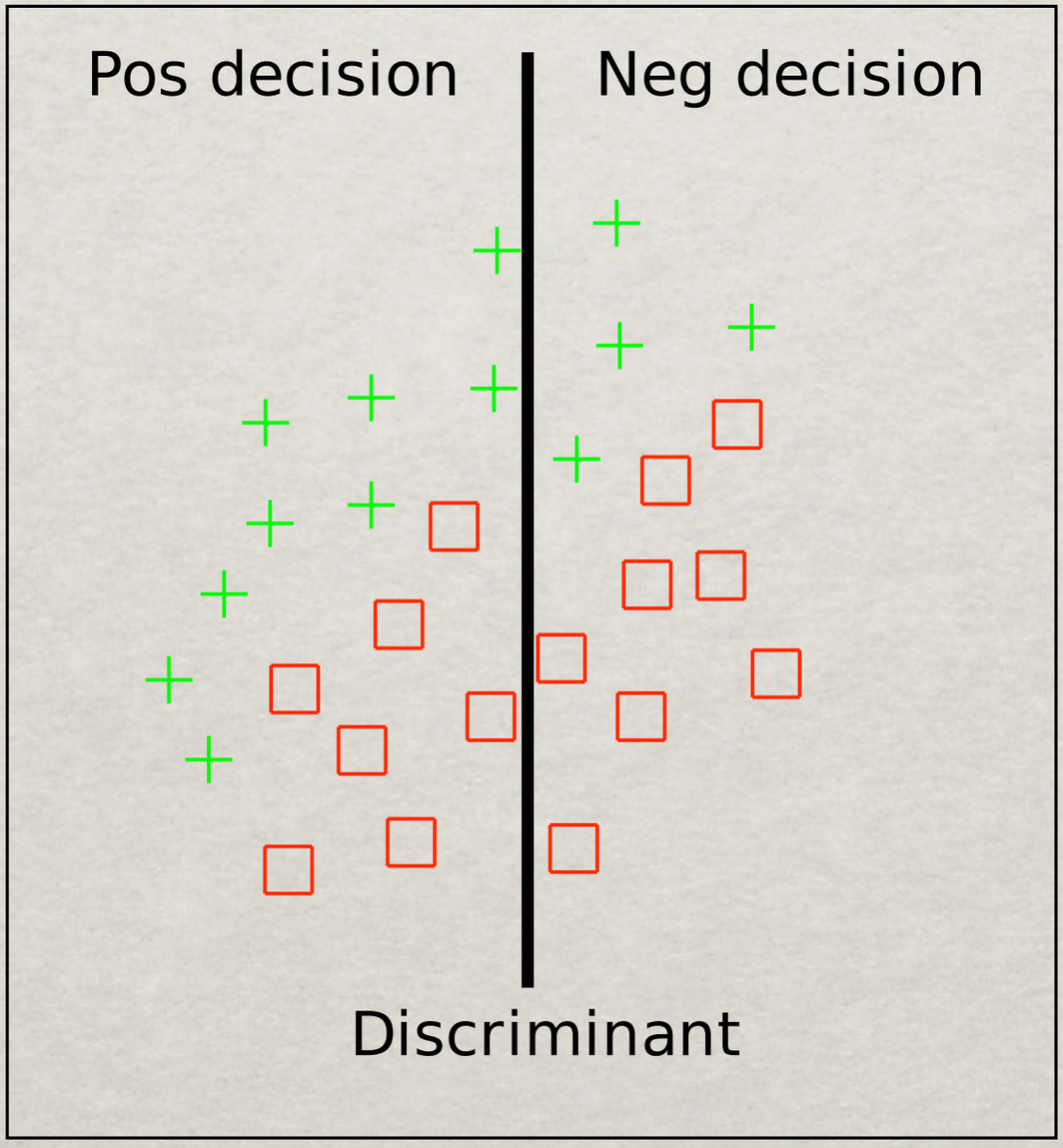
- Descriptor matching generates ordered *tentative correspondences*. When ground-truth is known, these can be evaluated with an *inlier frequency curve*, Chum&Matas, CVPR06.



- Good for RANSAC, and e.g. PROSAC (which uses the ranking).

ROC and PR curves

- Used for evaluating *binary classifiers* across a change of the discriminant.



ROC and PR curves

- Used for evaluating binary classifiers across a change of the discriminant.
- The optimal discriminant direction is often application independent, but the actual threshold is not.
- With ROC and PR curves, comparison can be done without committing to a specific discriminant.

ROC and PR curves

- Instead of a single performance measure we get a curve.

- Useful if criterion changes over time. E.g.

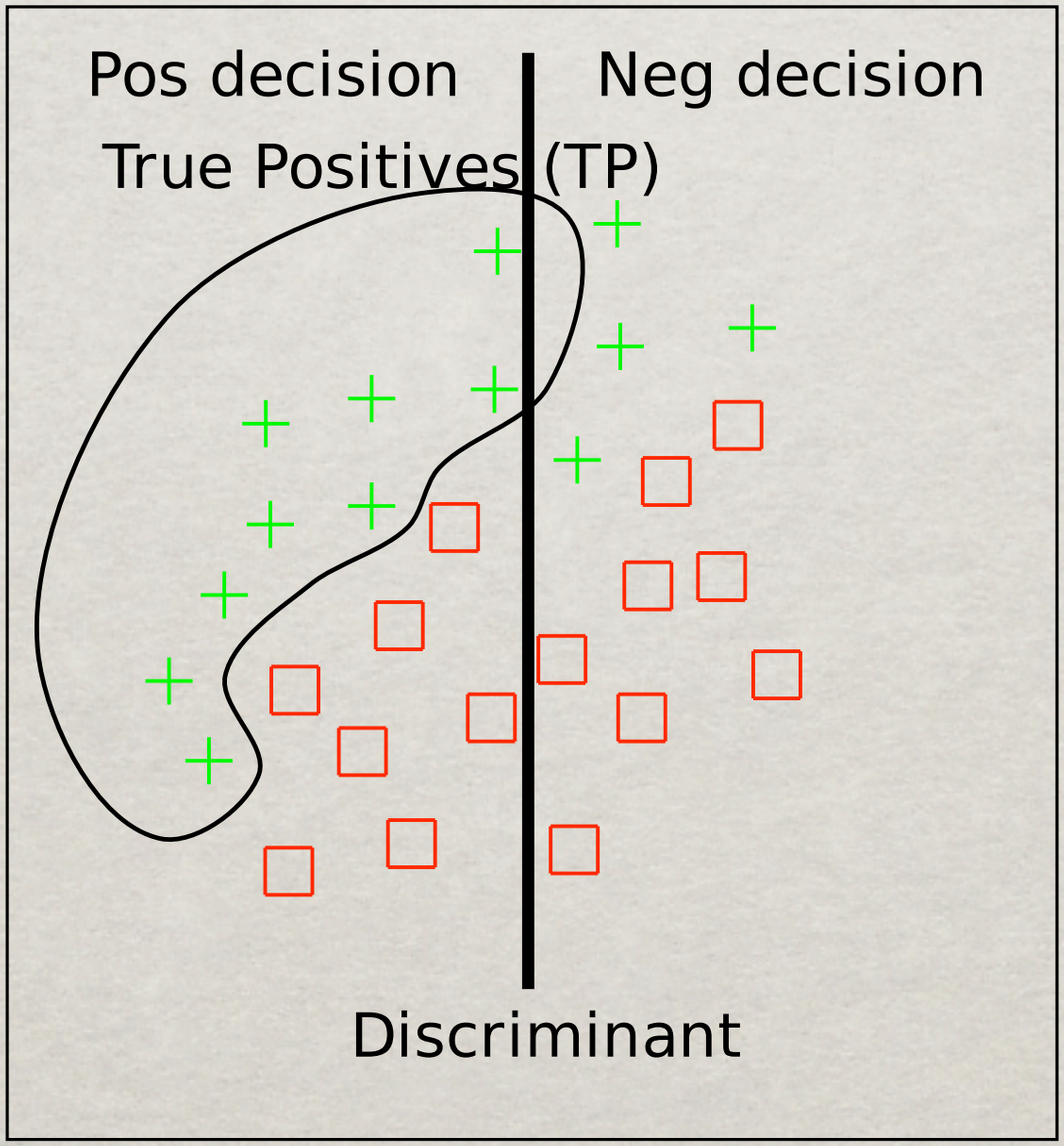
  1. Few false alarms might be most important.

  2. It might be very important not to miss a positive.

- To adapt, read curve in a different place.

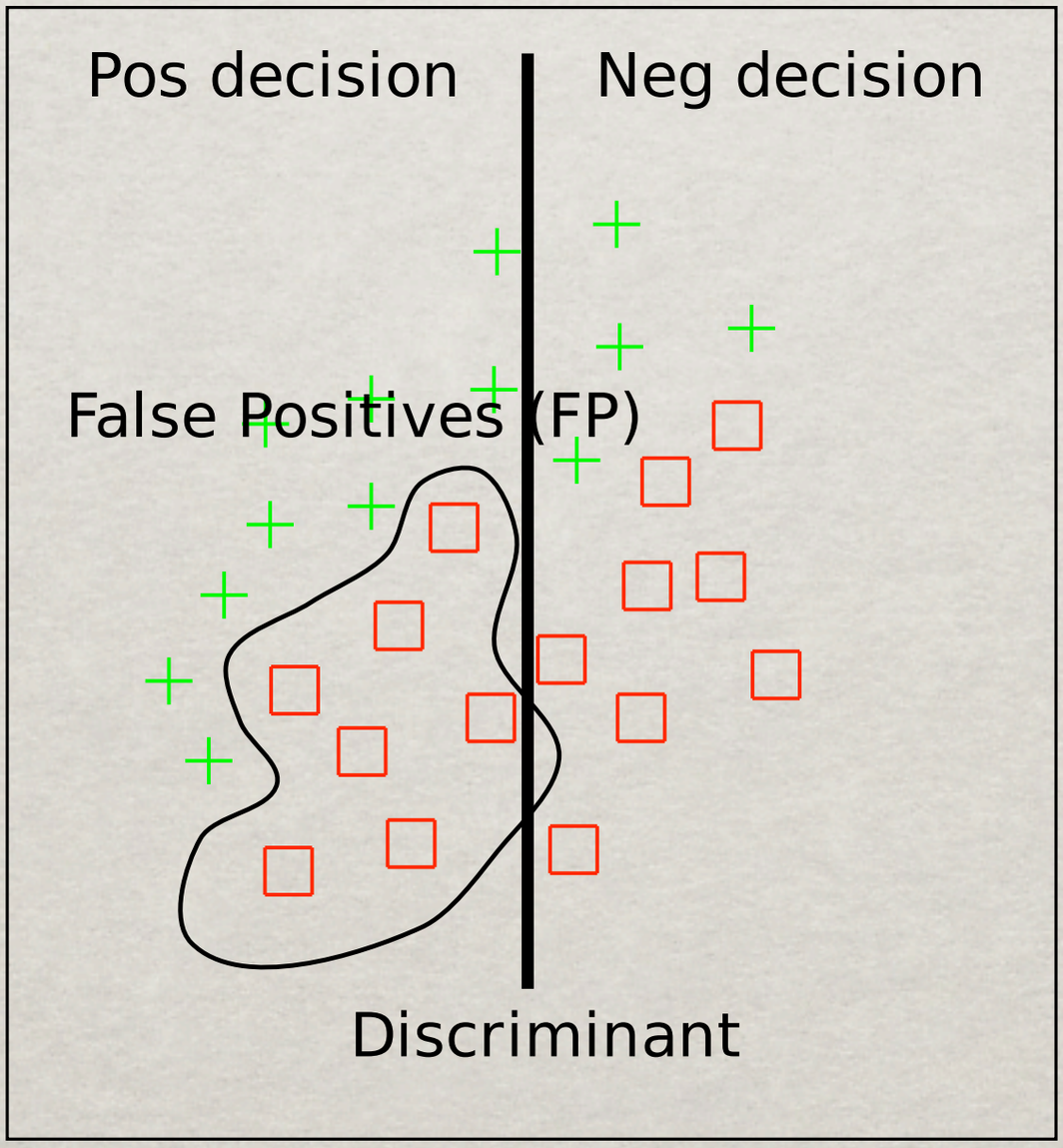
ROC and PR curves

- Used for evaluating binary classifiers.



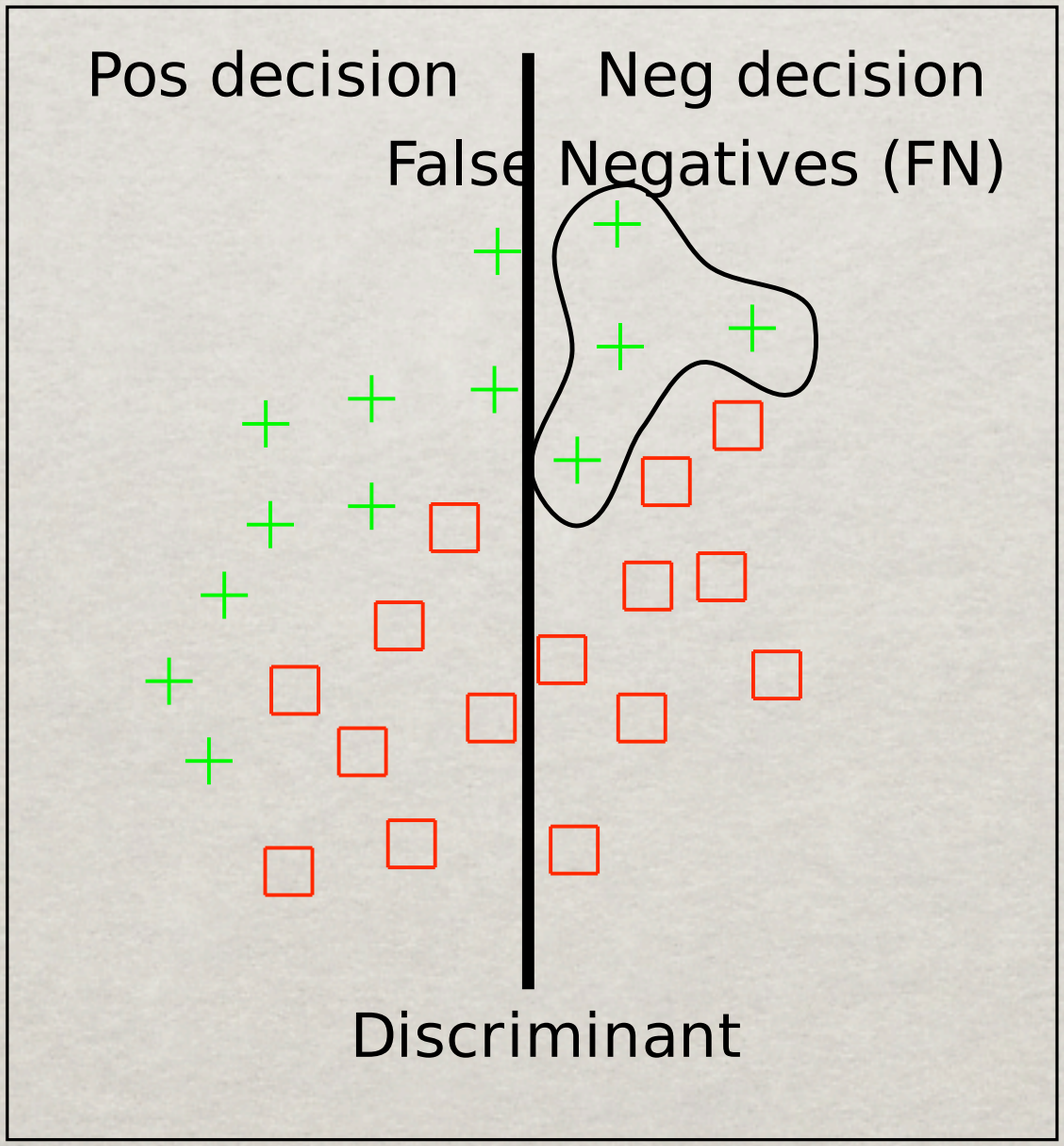
ROC and PR curves

- Used for evaluating binary classifiers.



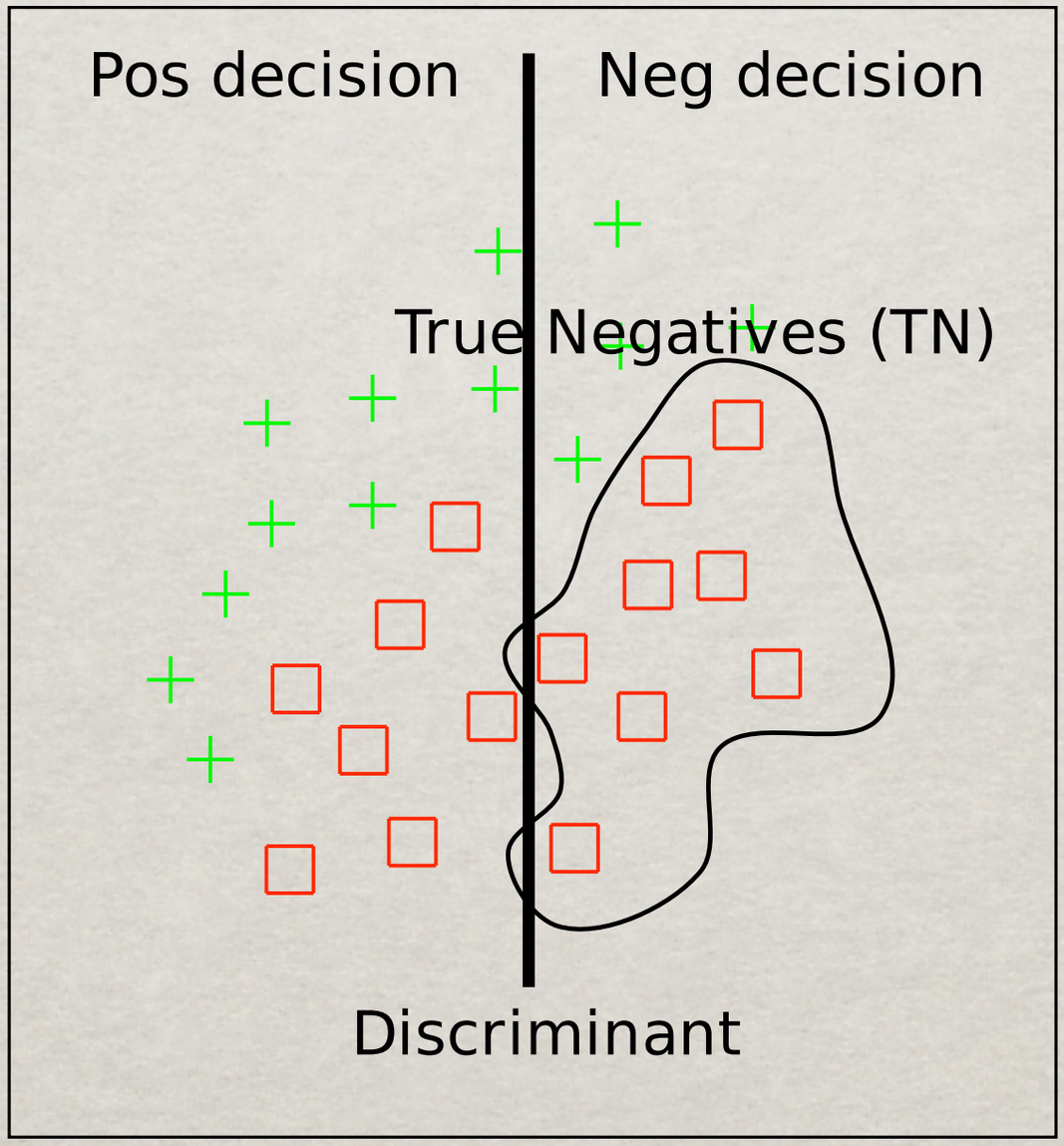
ROC and PR curves

- Used for evaluating binary classifiers.



ROC and PR curves

- Used for evaluating binary classifiers.



ROC curve

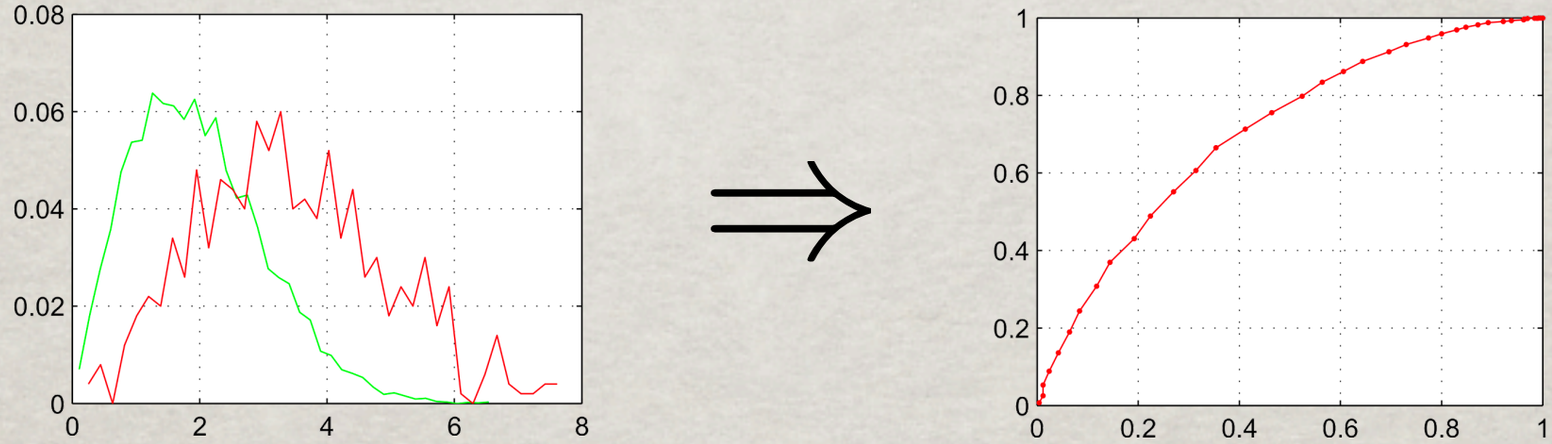
- Move discriminant, and plot True Postive Rate(TPR) against False Positive Rate(FPR)

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \qquad \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$

- Invariant to skewed datasets.
  Since normalisation is done with actual number of positives and negatives.

ROC from histograms

- ROC curves can used for evaluating matching performance as well. By using error histograms for inlier&outlier sets.



- Discriminant moving from left to right.

$$\texttt{TPR}(\epsilon) = \int_0^{\epsilon} p(\epsilon'|\texttt{inlier})d\epsilon' \qquad \texttt{FPR}(\epsilon) = \int_0^{\epsilon} p(\epsilon'|\texttt{outlier})d\epsilon'$$

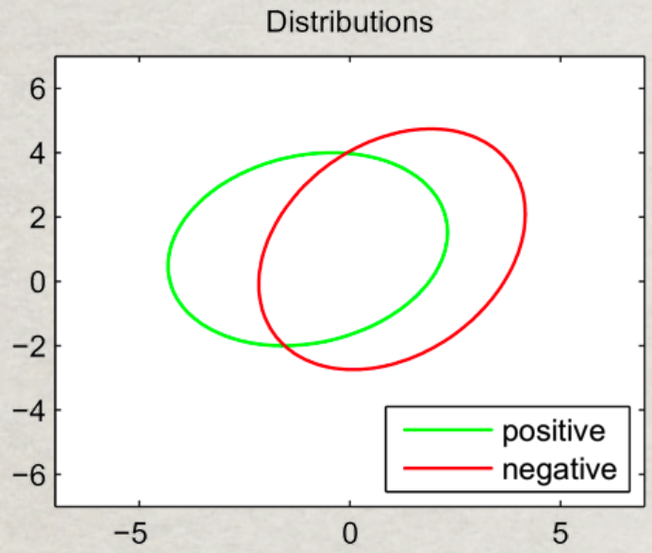
Precision-Reall curve

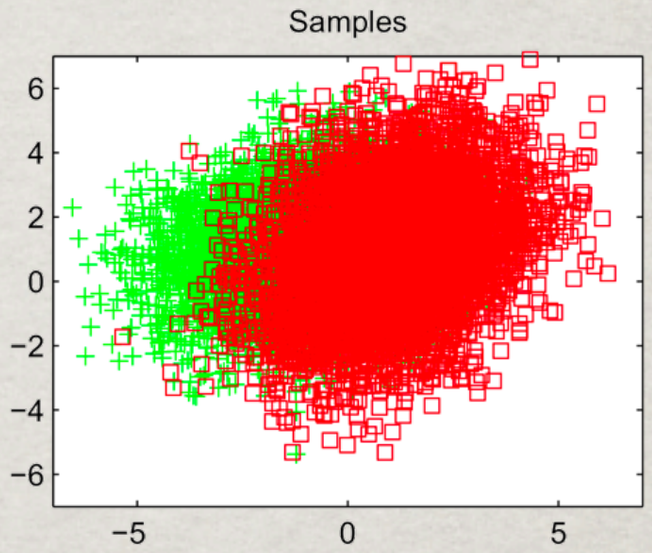
- Move discriminant, and plot Precision against Recall

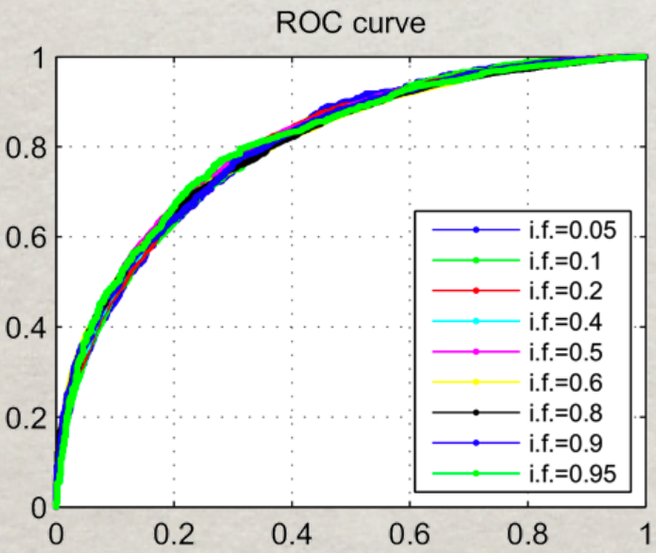
$$\texttt{Precision} = \frac{\texttt{TP}}{\texttt{TP} + \texttt{FP}} \qquad \texttt{Recall} = \texttt{TPR} = \frac{\texttt{TP}}{\texttt{TP} + \texttt{FN}}$$

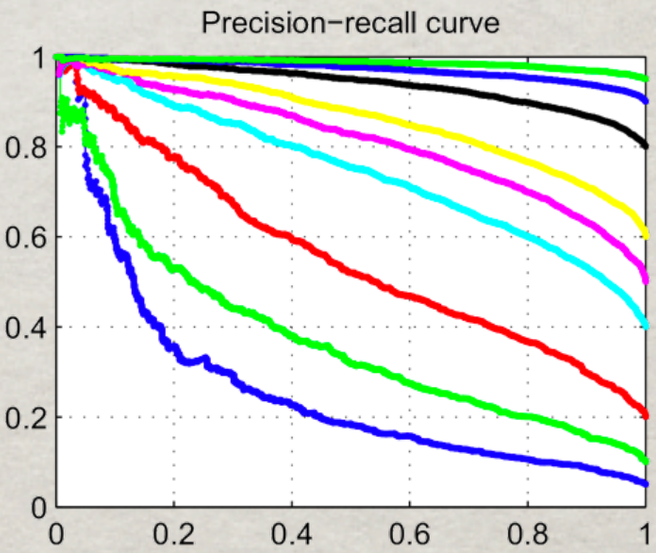
- Looks only at correctly reported positives.

ROC vs PR curves









- ROC ignores positive/negative sample ratio.

Exam

- Written exam contents:
  - The slides from all seven lectures
  - The six articles

Exam

- Time for the written exam: Nov 24, kl 13-14

Course Evaluation

- Webpage for course evaluation: http://cruz.isy.liu.se/wiki/index.php/VOR

- I will provide login and password.

Discussion

- Questions/comments on paper and lecture.