Object Detection
Lecture 10.1 - Descriptor-based detection

Idar Dyrdal

Recognition

Scene analysis and object recognition are still among the most challenging tasks in computer vision:

- Difficult to accurately name all possible objects in an image
- Objects may appear in different poses
- Cluttered scenes and occlusions
- Intrinsic variability within a class
- Non-rigid objects (people, animals)

Approaches:

- Object detection (trying to find objects of a specific type, i.e. faces, pedestrians, dogs etc.)
- Instance recognition (trying to find a specific object or individual, i.e. faces, rigid objects)
- Class recognition (Lecture 9.3)

UNIK4690

Object detection

Possible approach:

- Apply a recognition algorithm to every possible sub-window in the image (template matching)
- Generally slow and unreliable.

Better approach:

- Use special purpose detectors to rapidly find likely regions with objects of interest.
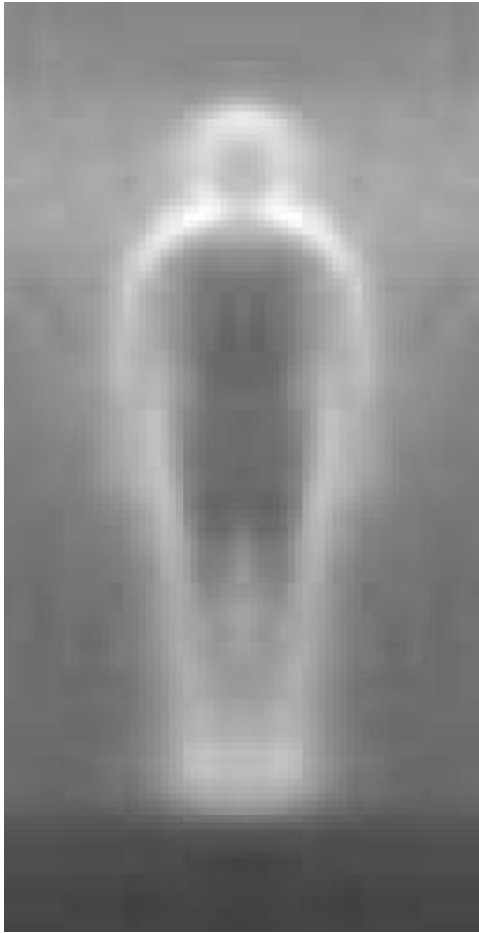
Topics:

- Face detection
- Pedestrian detection.

A: 57/57/1

(from Szeliski)

Pedestrian detection (example: Dalal-Triggs, 2005)

Navneet Dalal and Bill Triggs, Histograms of Oriented Gradients for Human Detection, CVPR05:
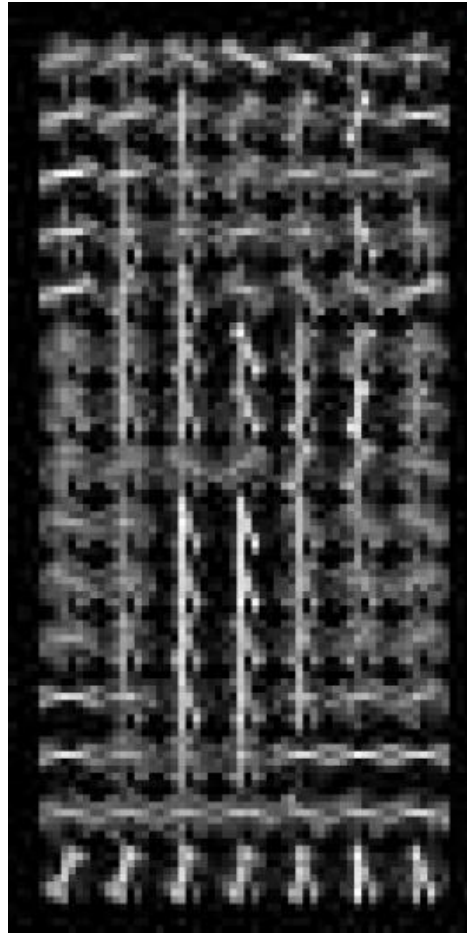
1. Extract windows of fixed size (64 x 128) at each position and scale
2. Compute histogram of gradient (HoG) features within each window
3. Compute a score for the window with a linear Support Vector Machine (SVM) classifier
4. Perform non-maximum suppression to remove overlapping detections with lower scores.
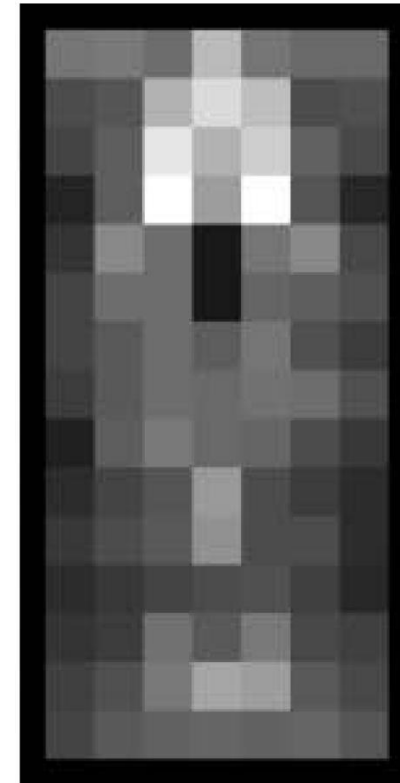
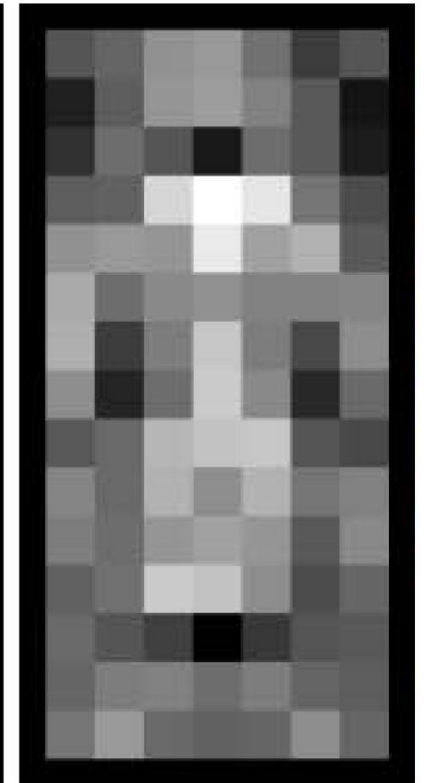# Pedestrian detection (2)

Gradient image
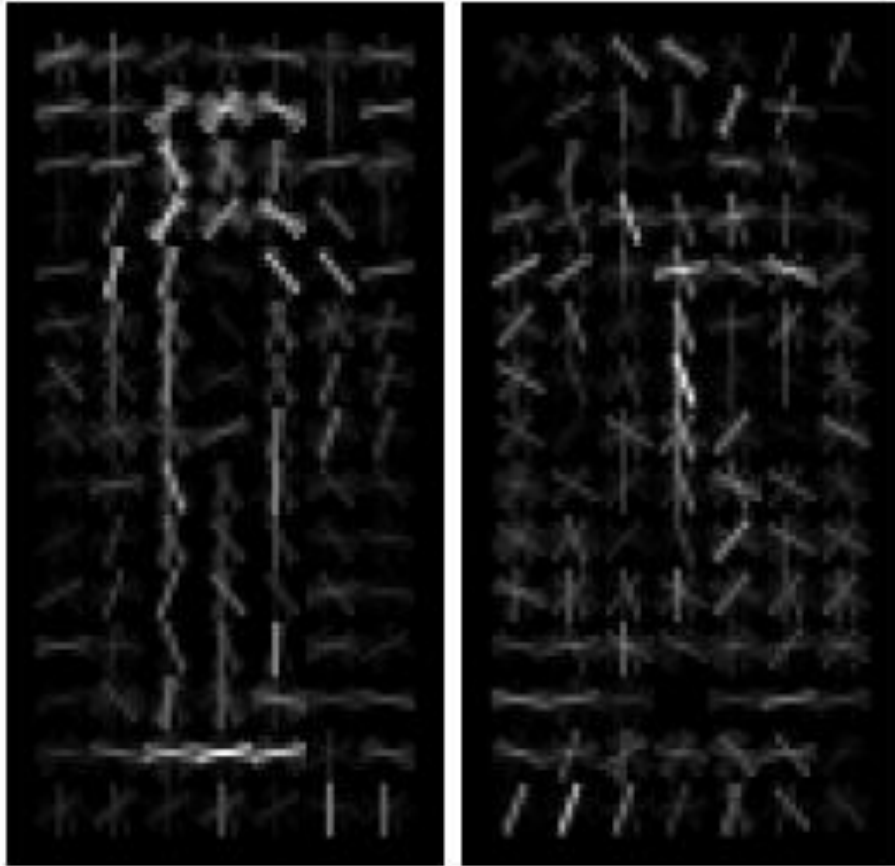
HoG (weighted and interpolated)

Positive weights   Negative weights

SVM weights

Pedestrian detection (3)

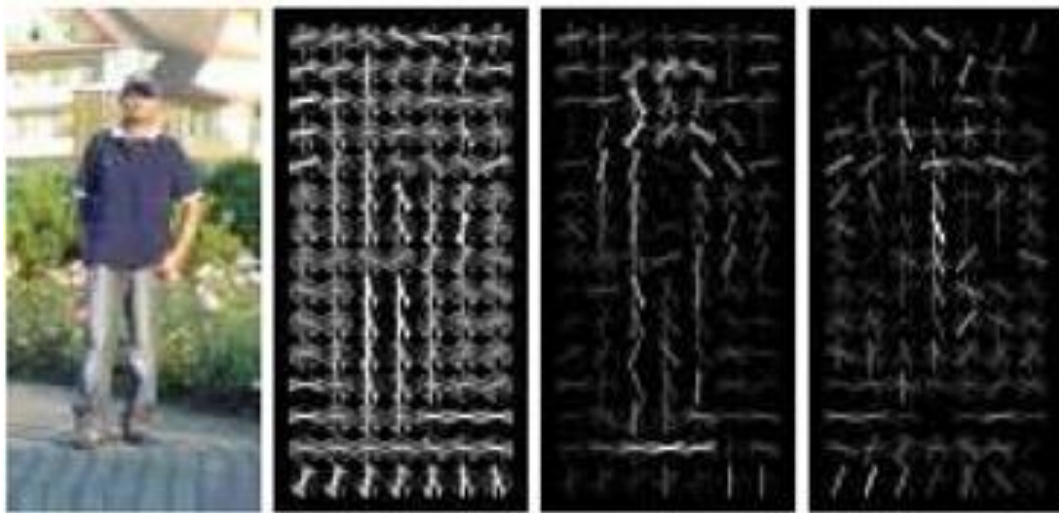HoG weighted with SVM-weights

Positive weights   Negative weights
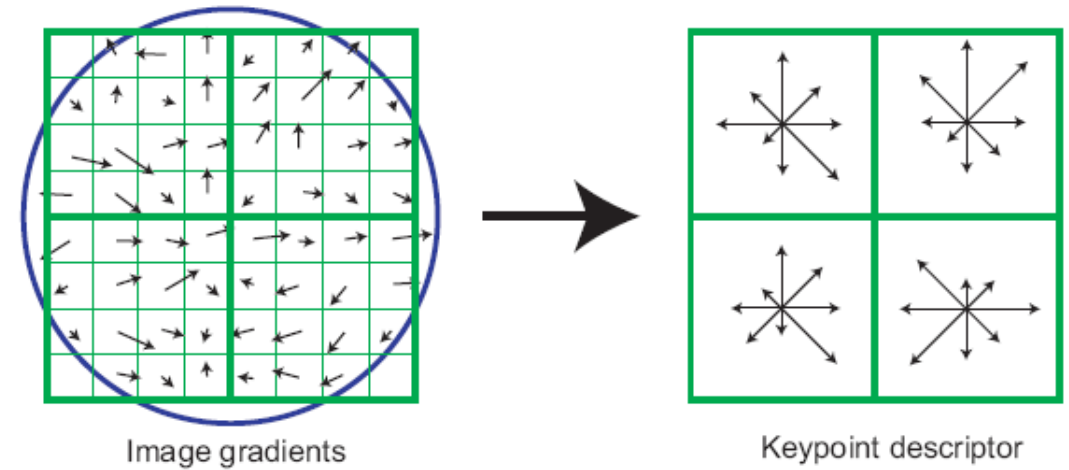
$$w^t x - b$$

Score = 0.16 > 0 => «Pedestrian»
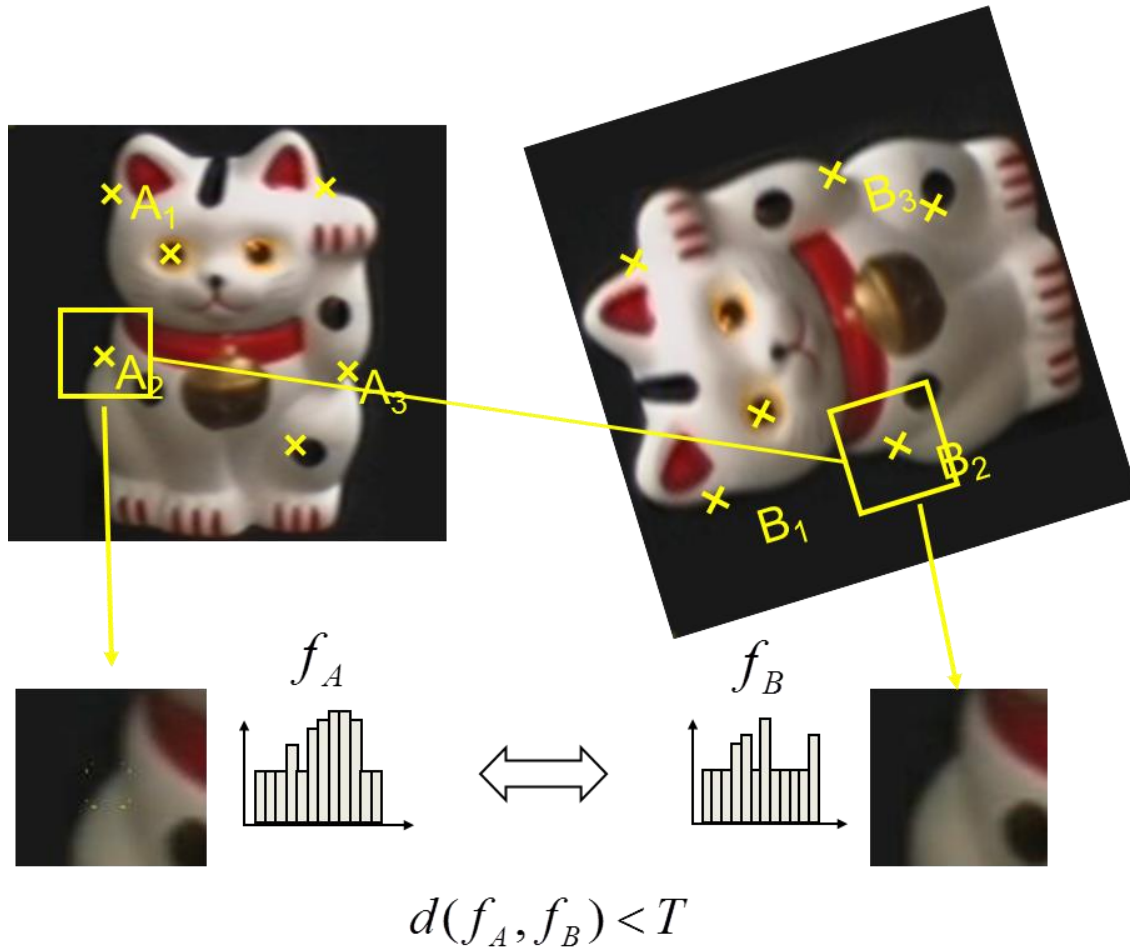
Feature descriptors

HOG (Histogram of Gradients)

Image gradients → Keypoint descriptor

SIFT (Scale Invariant Feature Transform)

Feature descriptors (2)

- HoG descriptors

- SIFT

- SURF

- Binary descriptors
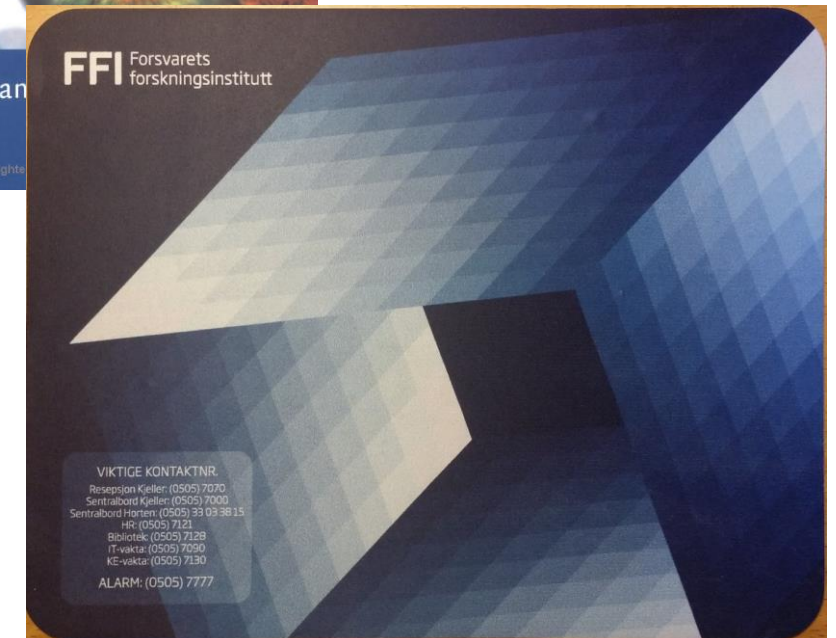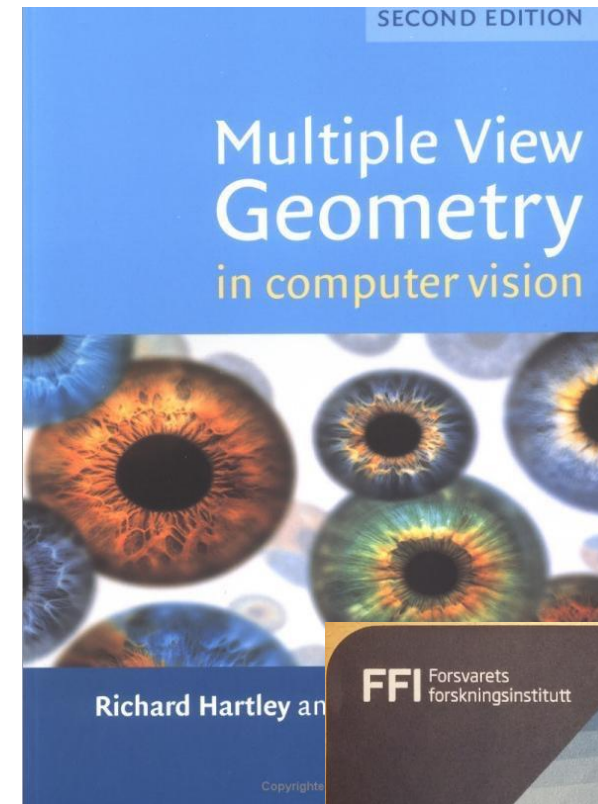
$$f_A \qquad f_B$$

$$d(f_A, f_B) < T$$

Instance Recognition

Problem:

- Re-recognition of known objects (2D or 3D)
- Rigid objects viewed from novel viewpoints
- Cluttered background
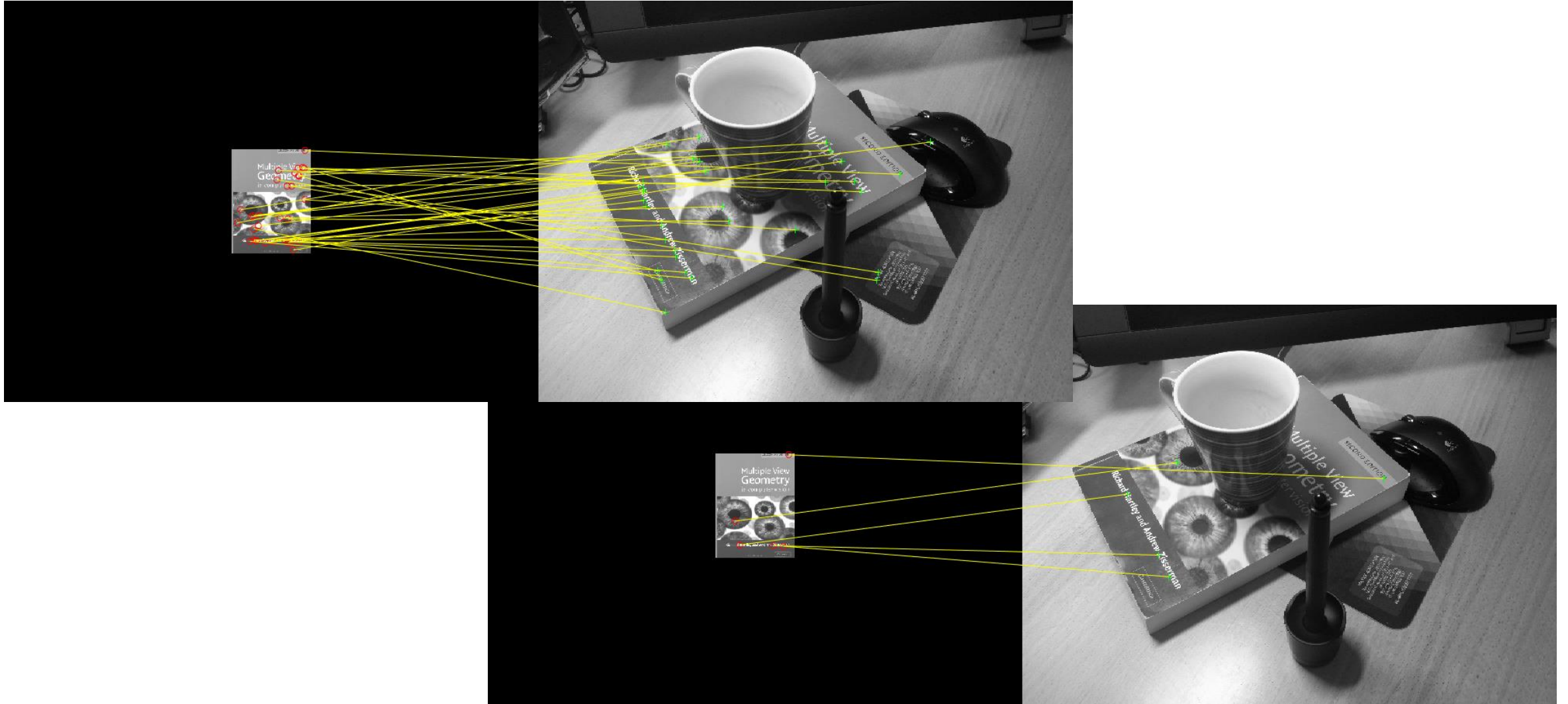- Partial occlusions.

Approach:

- Extract informative 2D features from new images and match to corresponding features (descriptors) for objects in the database
- Find geometric transformation aligning the two sets of features.
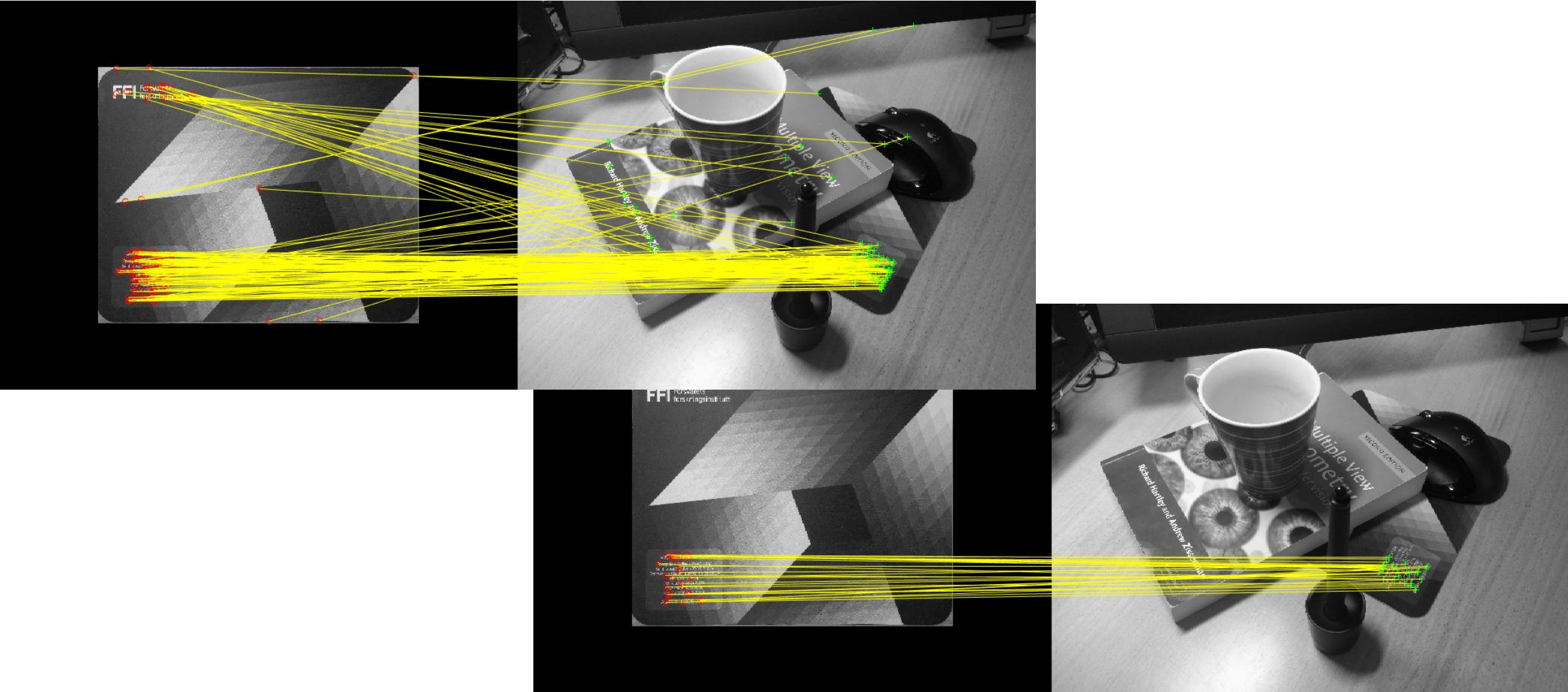
Scene with clutter and partial occlusions

Descriptor matching

Descriptor matching (2)

Geometric alignment

Summary

Descriptor-based detection:

- Feature Descriptors
- Object Detection
- Instance Recognition

More information: Szeliski 14.1 and 14.3