Digital Image Processing

Shape Matching

Deformable Template Matching

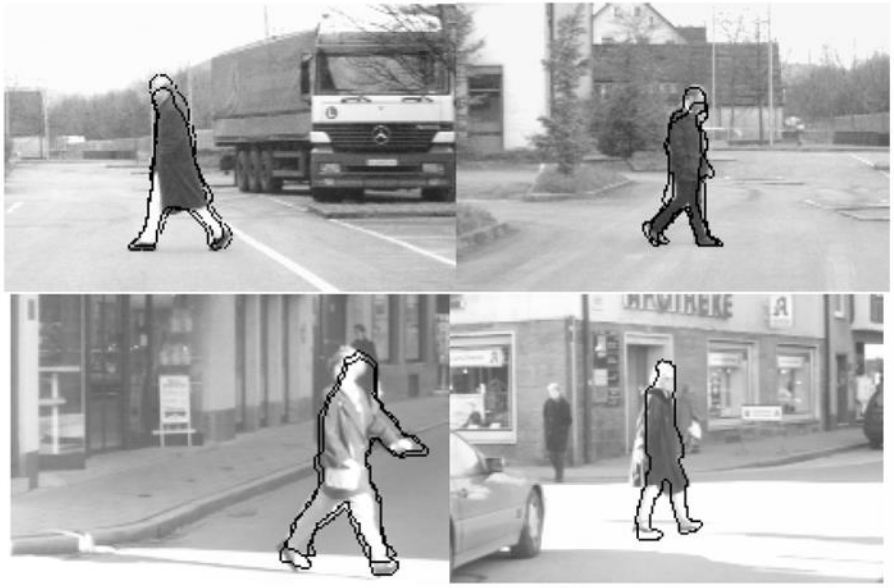
Scale Invariant Feature Transform

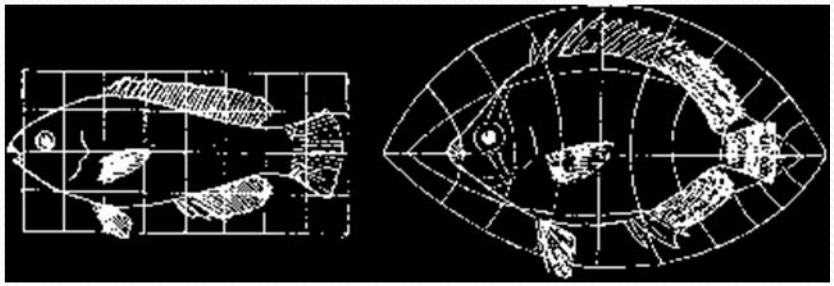
Topics

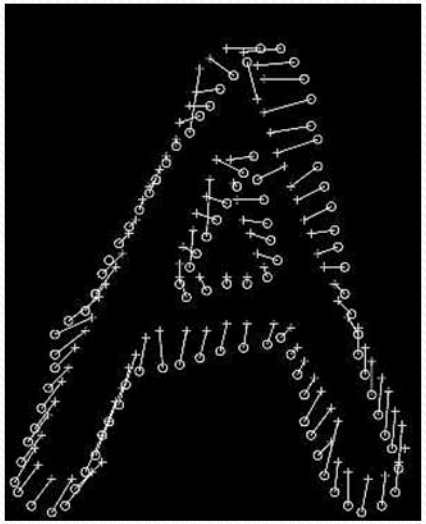
- Object Detection
  - Shapes
  - Shape Detection using Energy Minimization
- Scale
  - Hierarchical Representation of Images
  - Object Features
  - Scale Invariance
  - Scale Invariant Feature Transform (SIFT)

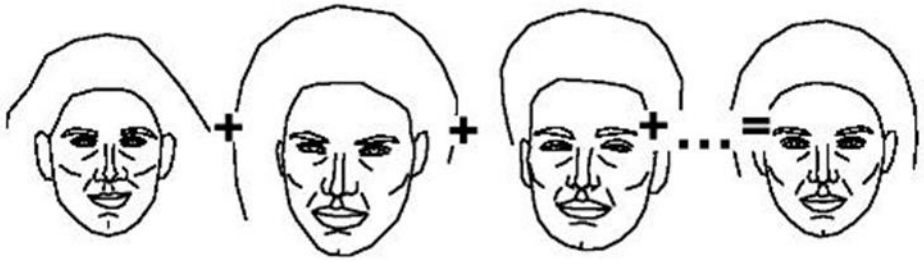
Motivation

- What is a shape?









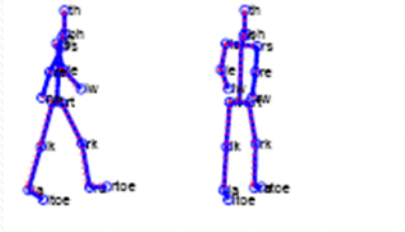
Shape

- Defn 1: A set of points that collectively represent the object
  - We are interested in their location information alone!!

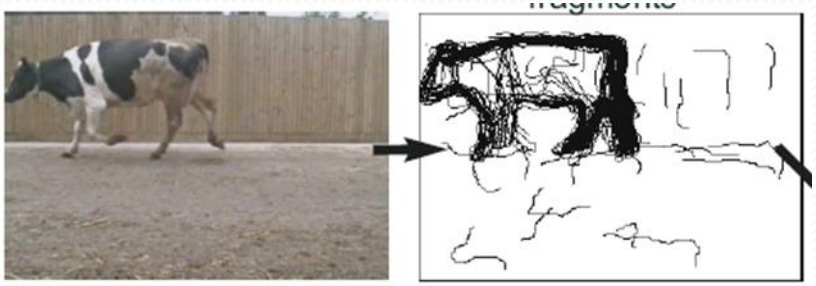
- Defn 2: Mathematically, shape is an equivalence class under a group of transformations
  - Given a set of points X representing an object O, and a set of transformations T, shape S={t(X)| t\in T}

Applications of Shapes

- Recognition and Detection



Pose



Shape Detection/Recognition Issues

- Representation
  - Moments
  - Fourier descriptors
  - Chain Code
  - Etc.
- Matching
  - How to compute distance between shapes?
- Challenges in recognition
  - Information loss in 3D to 2D projection
  - Articulations
  - Occlusion….
  - Scale, Orientation, Rotation Invariance

How to Compare Shapes?

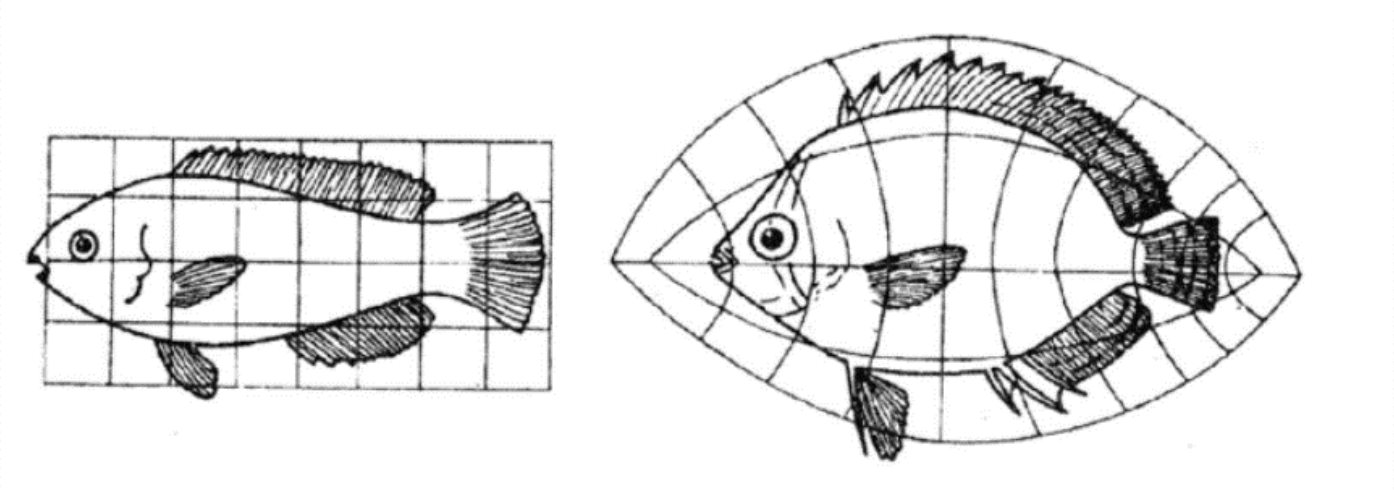
Fig. 1. Examples of two handwritten digits. In terms of pixel-to-pixel comparisons, these two images are quite different, but to the human observer, the shapes appear to be similar.

How to Compare Shapes?

- Hausdorff distance
  - Hausdorff distance is the *maximum distance of a set to the nearest point in the other set*
  - *It find the upper bound on the distance of any point from A to a point in B*

Invariance/ Robustness

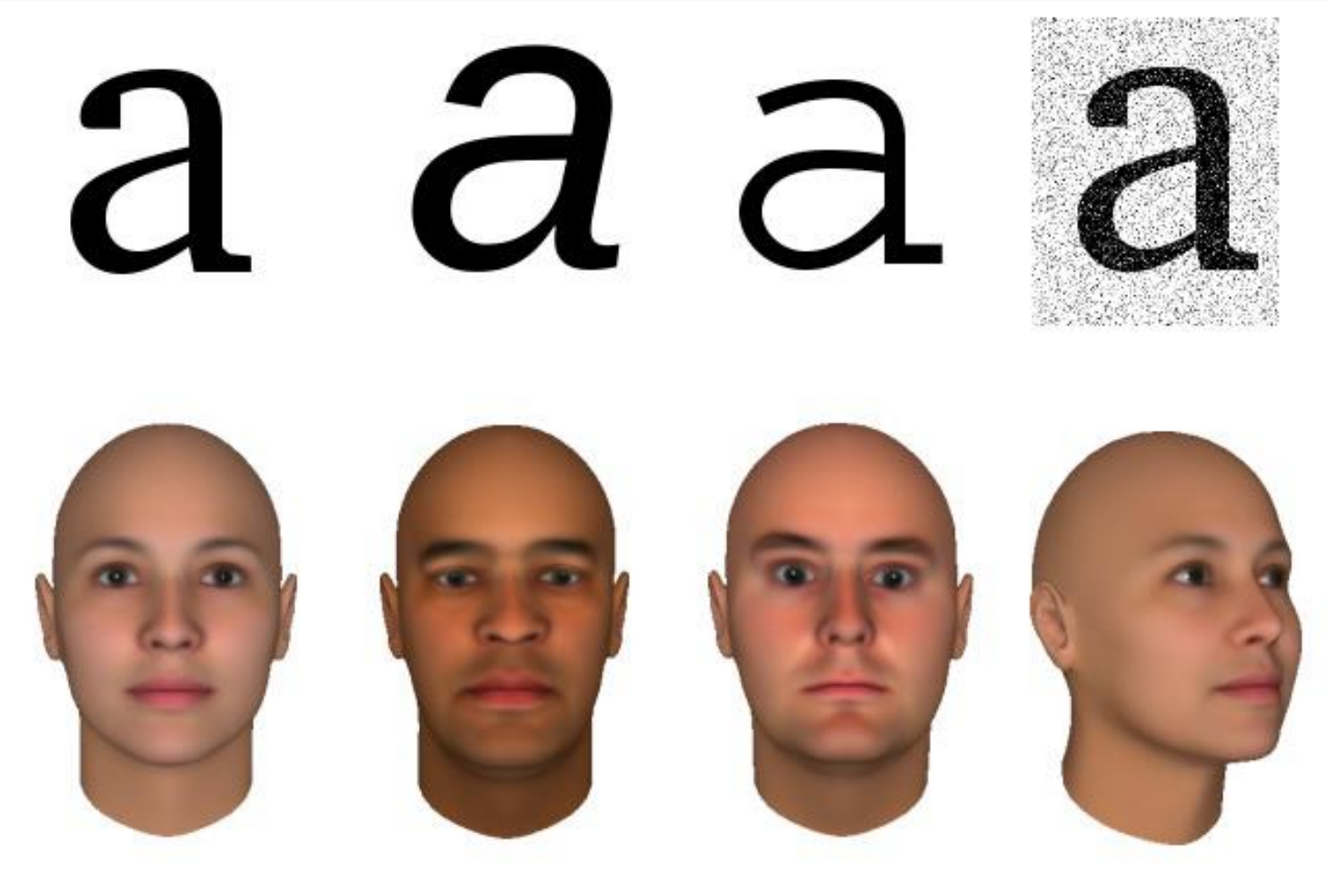
- Translation
- Scaling
- Rotation



Template Matching

- **Template:** Something formed after a model or prototype, a copy; a likeness.
- **Matching:** To compare in respect of similarity; to examine the likeness or difference of.

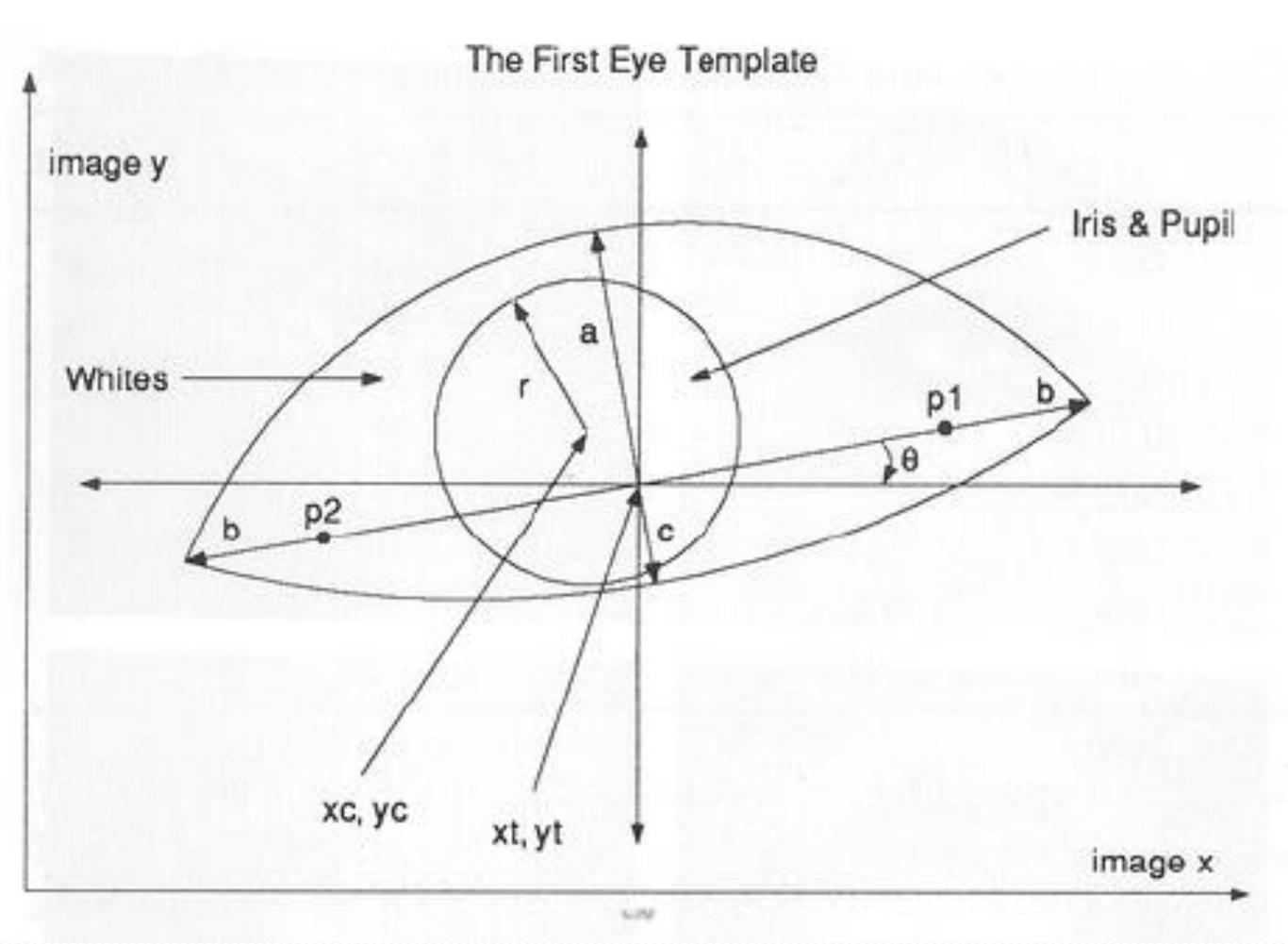
Template Matching



Deformable Models

- A prototype model which describes the shape of a class of objects (generally a parametric model)
- A set of parameters to deform/change the model before matching with the object(s)
- A metric to decide when an object is detected

Example: Human Eye Detection



Eye Template Description

- The circle representing the iris, characterized by its radius r and its center xc . The interior of the circle is attracted to the low intensity values while its boundary is attracted to edges in image intensity

Matching Metric

- An energy function is defined to determine where a matching occurs
- The energy function has terms corresponding to:
  - Eye boundaries
  - Iris boundaries
  - Intensity of iris pixels
  - Intensity of white pixels around iris

Object Matching based on Invariants

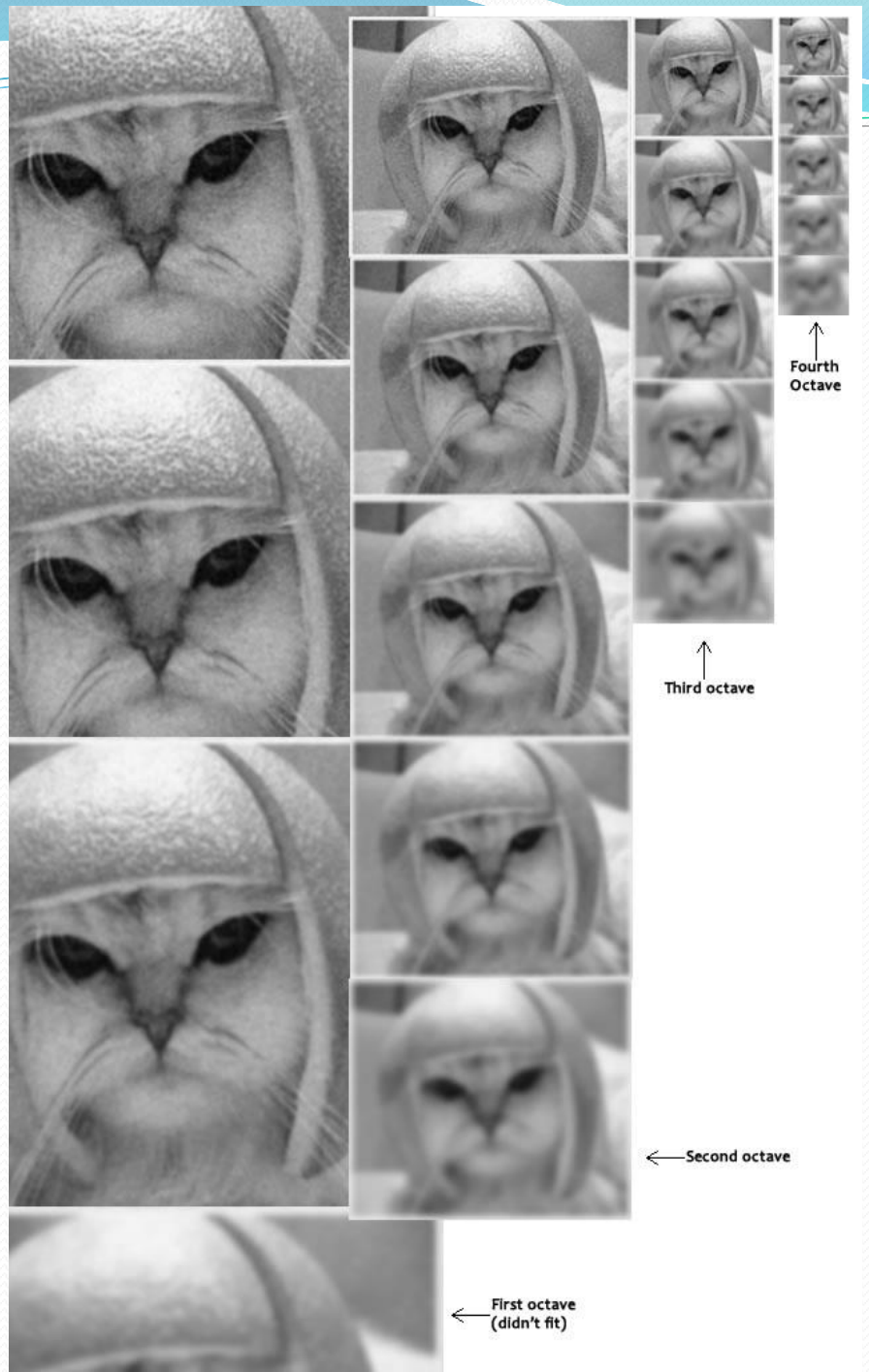
- Many object detection methods are sensitive to scale or rotation
- The goal here developing a matching method which is insensitive to scale and rotation

Scale Invariant Feature Transform

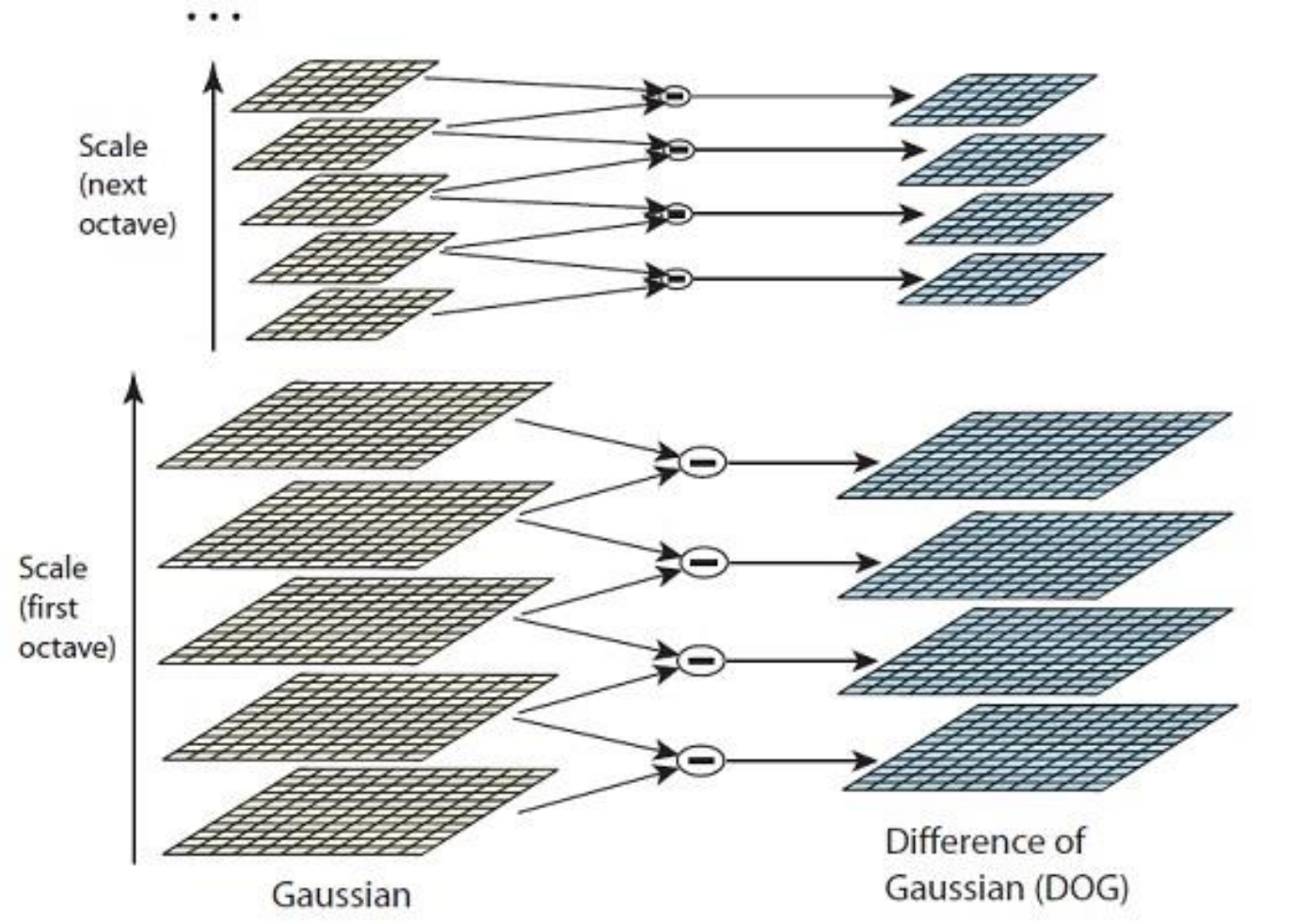
- The method consists of the following stages:
  - Constructing a scale space
  - LoG Approximation
  - Finding key points
  - Eliminating bad key points
  - Assigning an orientation to the key points
  - Generating SIFT features

Constructing a scale space

- Generates progressively blurred out images with different scales

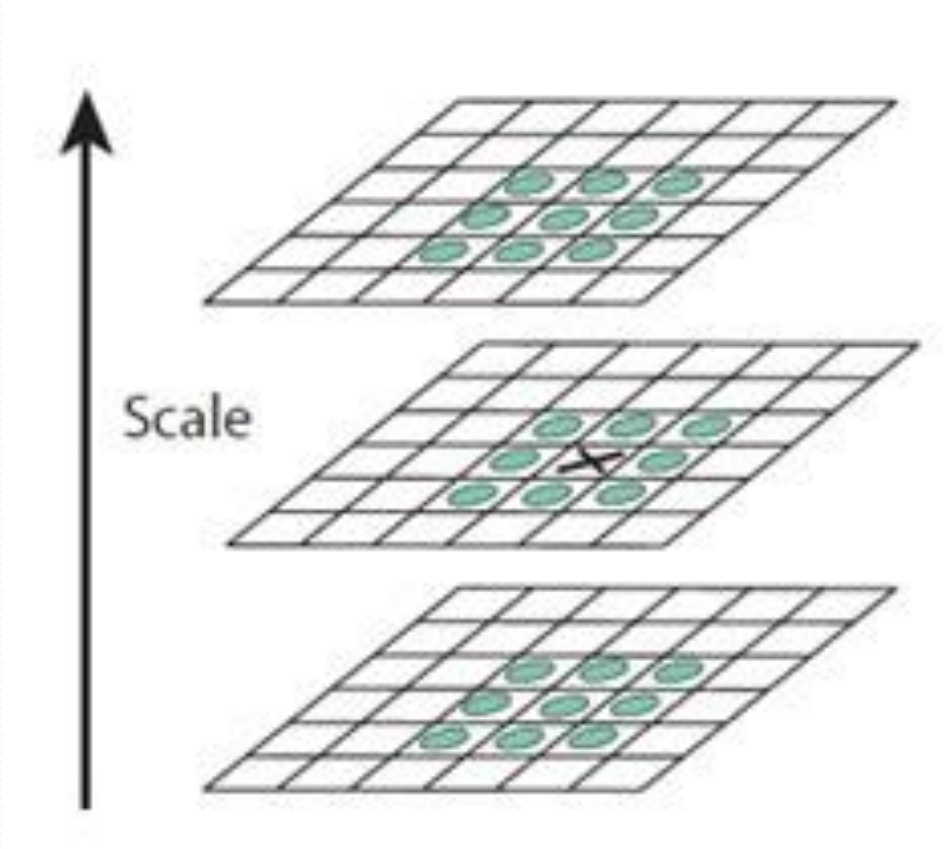

Fourth Octave
Third octave
Second octave
First octave (didn't fit)

LoG Approximation



Finding key points

- Locate the maxima and minima



Eliminating bad key points

- Removing low contrast features
- Removing edges

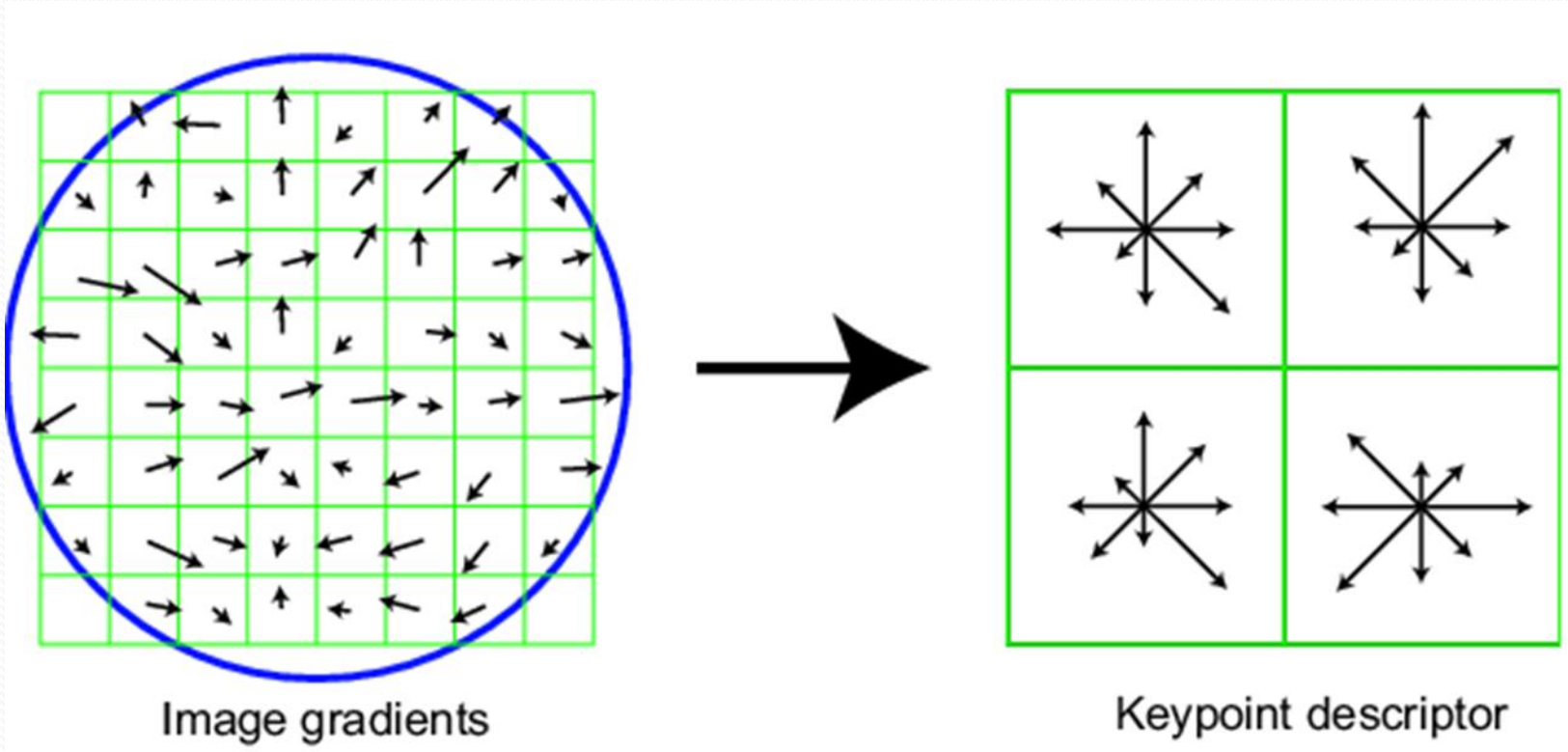
Key Point Orientation

- The magnitude and orientation is calculated for all pixels around each key point. Then, A histogram is created for this.
- The orientation can be quantized
- The peaks of the histogram define the orientation of the key point

Create Signatures

- A 16x16 window is defined around the key point. The window is divided into sixteen 4x4 windows
- Create array of orientation histograms
- 8 orientations x 4x4 histogram array = 128 dimensions

Create Signatures



Create Signatures

- Improvements:
  - Orientation invariance: To achieve rotation independence, the key point's rotation is subtracted from each orientation
- Larger cell (6-8 pixels) can be used for better approximation.
- Fine orientation binning will give better approximation for orientation
  - 9 bins/180^{O} vs. 8 bins/360^{O}

Questions?