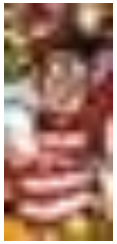
Pattern Recognition, Visual search and Hough Trasform

Where's Waldo?

Scene



Template

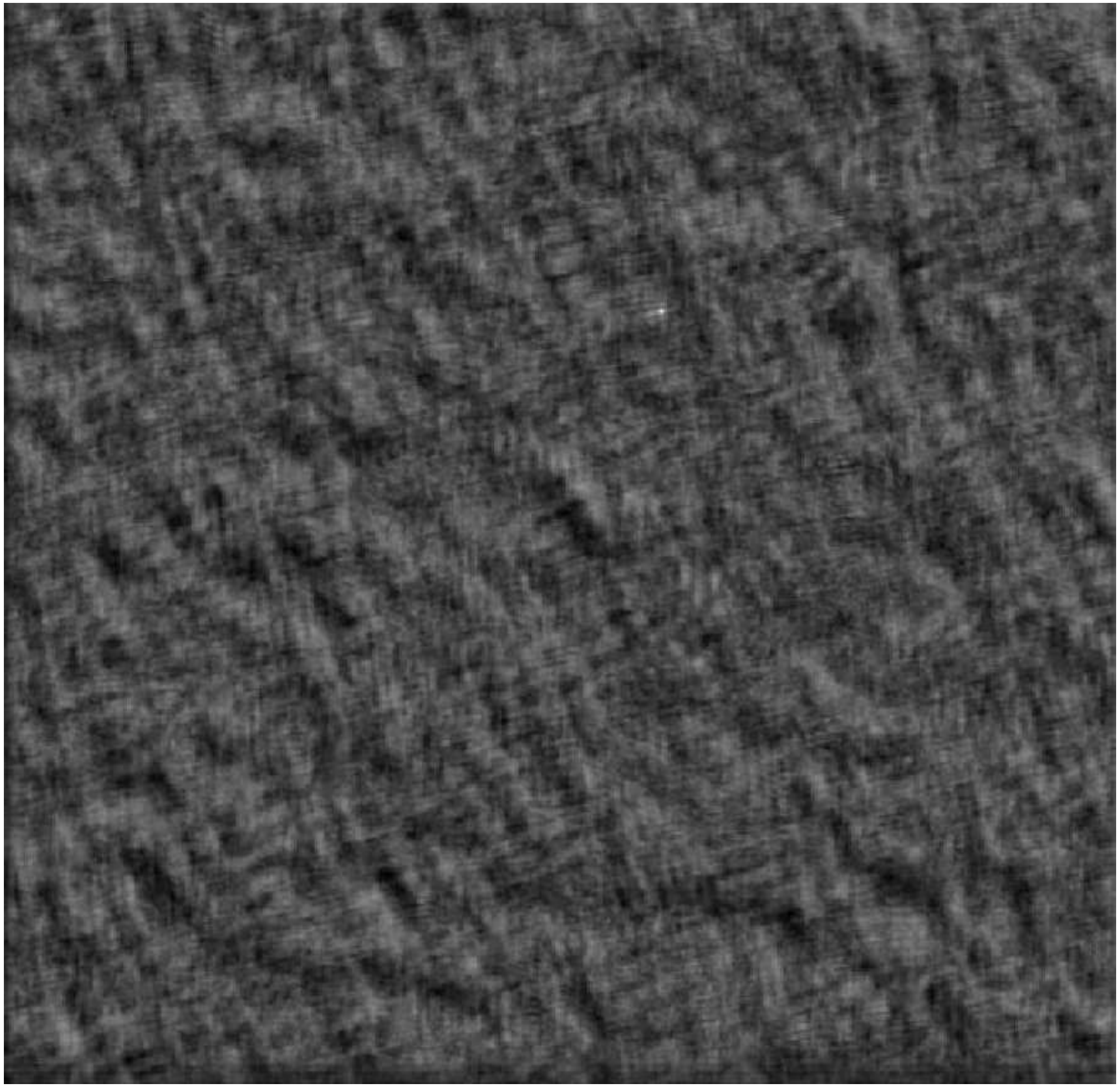
Where's Waldo?

- A simple idea:
  - I move the target on the image and I compare target and image minimizing an error function

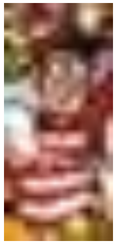
$$E(y,x)=\sum_{i,j}\left(I(y+i,x+j)-T(i,j)\right)^2$$

$$E(y,x)=\sum_{i,j}\left|I(y+i,x+j)-T(i,j)\right|$$

Where's Waldo?

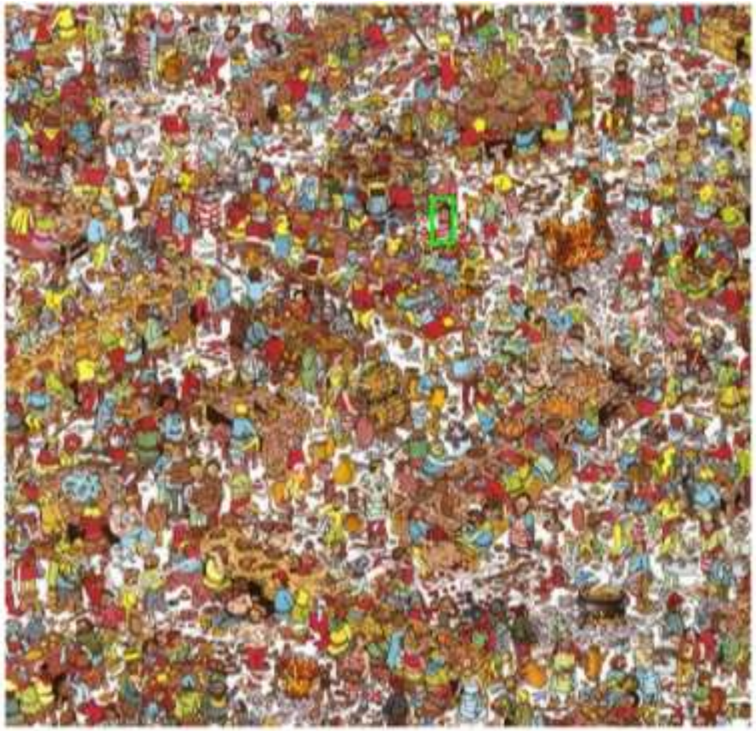


Scene

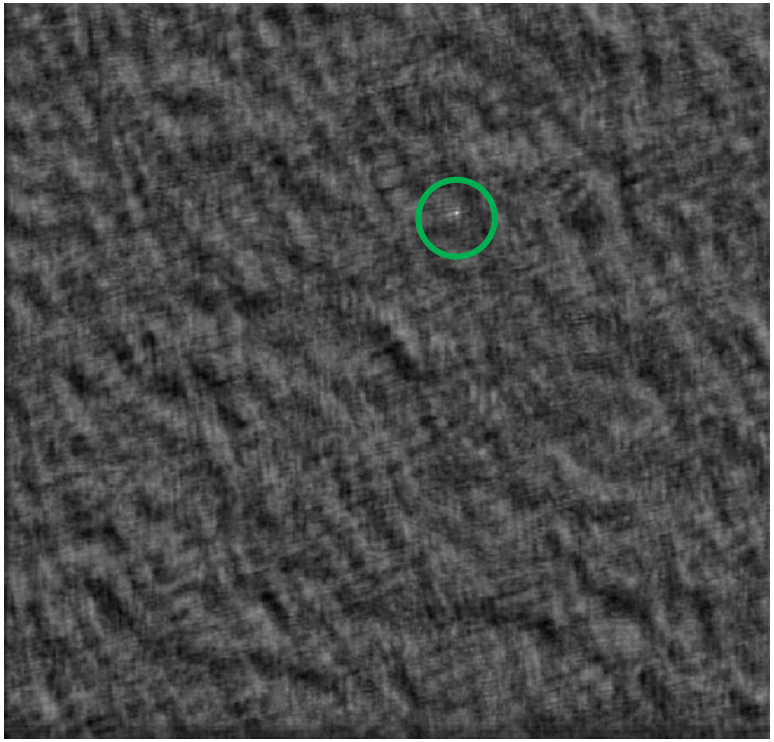


Template

Where's Waldo?



Detected template



Correlation map

PR and artificial visual search

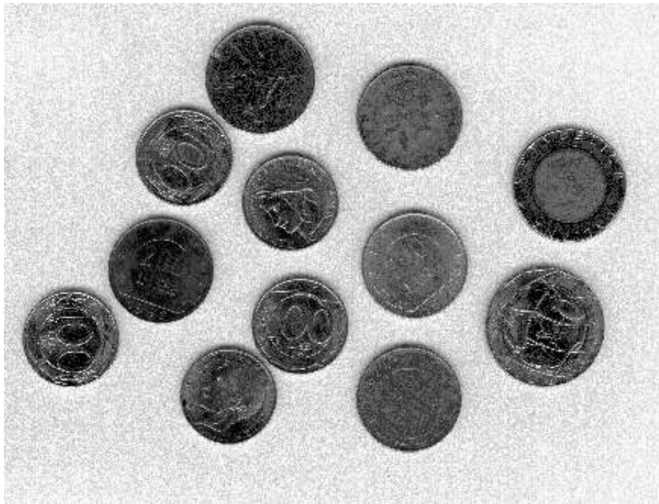
- Given an unknown input image segment and the basic properties (template) of the object-target, the problem is to determine if the segment belongs to the target class

- Difficulties can arise when large variations and distortions are expected in this segment. In general this is the case because objects appears:
  - roto-translated and with a scaling factor



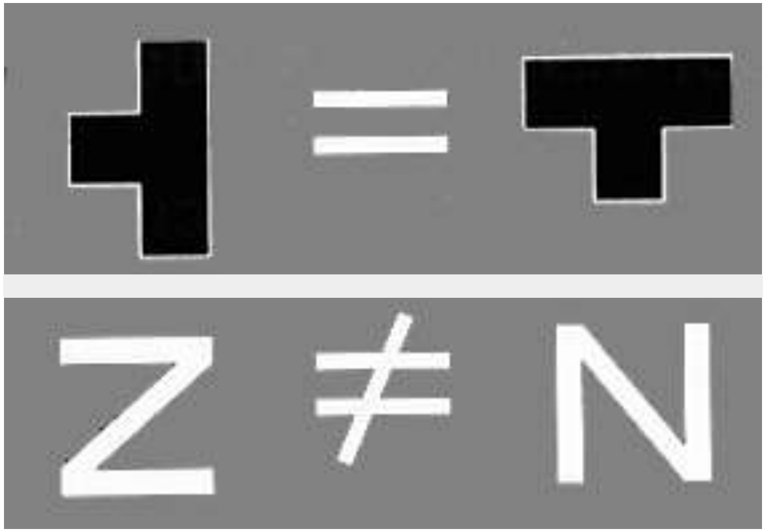
PR and artificial visual search

- In general this is the case because objects appears:
  - with shading, luminance and color changes
  - overlapped, occluded and noisy
  - rigid, semi rigid or even flexible



Model definition

- The real world contains, in general, high variability and variety levels for mathematical and statistical models to describe the model of the class.
- The components outside such descriptions are commonly termed noise. An automatic system – and maybe the human mind itself – is necessarily endowed with models to interpret reality – where the so called **context** is part of.
- In computer vision ***context*** can be described through a particular configuration of internal parameters and pragmatically a context is valid if the automatic interpretations, in the current scenario, correspond to an acceptable extent, to the target goals.



A taxonomy of PR approaches

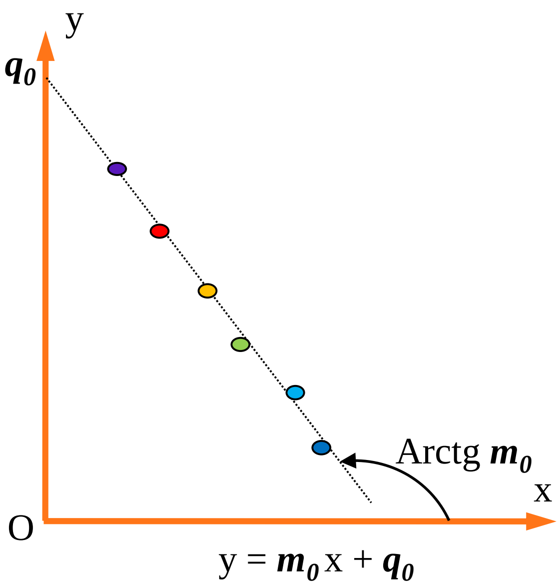
- Generally we can divide the different objects recognition techniques in:
  - Appearance-based methods, in which are used example images (called **templates**) of the objects to perform recognition; problems rise because objects look different under varying conditions:
    - ✓ Changes in lighting or color;
    - ✓ Changes in viewing direction;
    - ✓ Changes in size or shape.

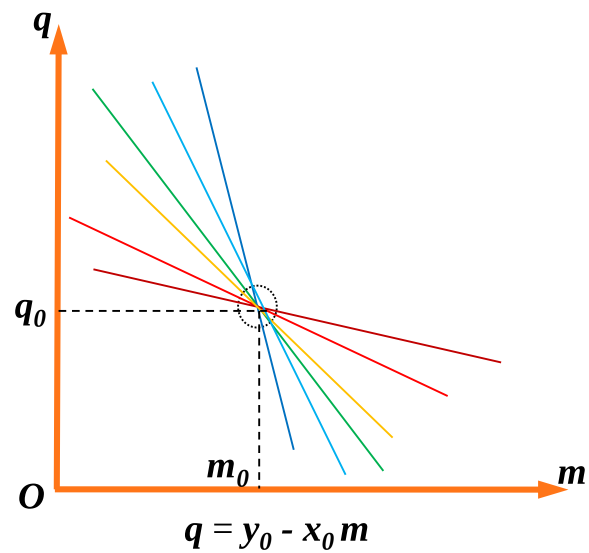
    Techniques: Edge Matching, Divide-and-Conquer Search, Greyscale Matching Edges, Gradient Matching, …
  - Feature-based methods, a search is used to find feasible matches between object features and image features. There are different solutions used to extract features from the objects to be recognized and the images to be searched such as:
    - ✓ Surface patches;
    - ✓ Corners;
    - ✓ Linear edges.

    Techniques: Interpretation Trees, Hypothesize and Test, Scale-Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), HOG - Histogram of Orientated Gradients, …

Hough Transform

- The Hough transform has been introduced in 1962 by Paul Hough for the detection of straight lines.
- Each contour point identified in an image can support the existence of the set of straight lines crossing its location. If a straight line is present in the image, and N of its points are detected, N sets of lines receive a contribution but only the common single straight line receives N contributions.





Hough Transform

- In general an analytical function in a two dimensional space (image space) is defined by a small set of parameters
- An equation described the relation between the coordinate and the parameters.
  - $f((x,y), (a_1,a_2,\ldots,a_n))=0$
  - (x,y) is a point in the image space (IS)
  - $(a_1,a_2,\ldots,a_n)$ is a set of values in the parameter space (PS)
- A point in the PS defines exactly an analitical function

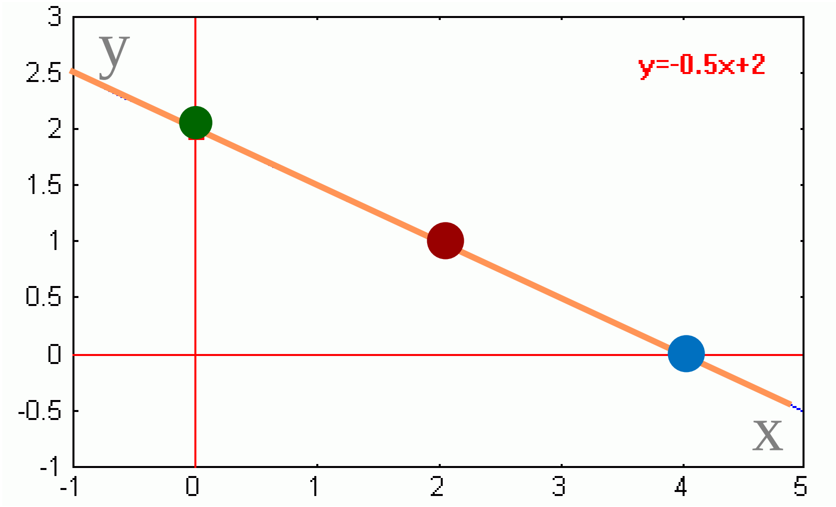
HT: searching straight lines

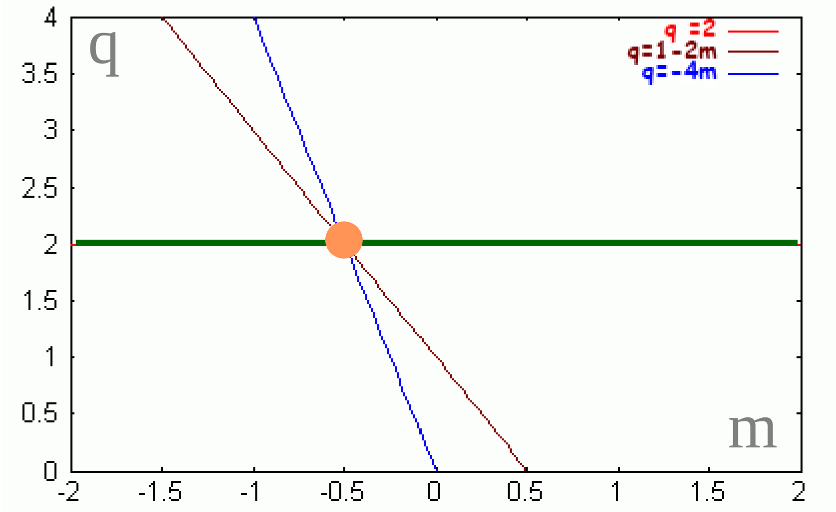
- Classical stragth line equation:

$$y = mx+q$$

$$f((x,y), (m,q)) = y-mx-q = 0$$

- Given a point (x_i,y_i) in the image space (IS) the equation $q = y_i - mx_i$ describes the locus of points of the parameter space (PS) representing the set of straight line crossing (x_i,y_i)





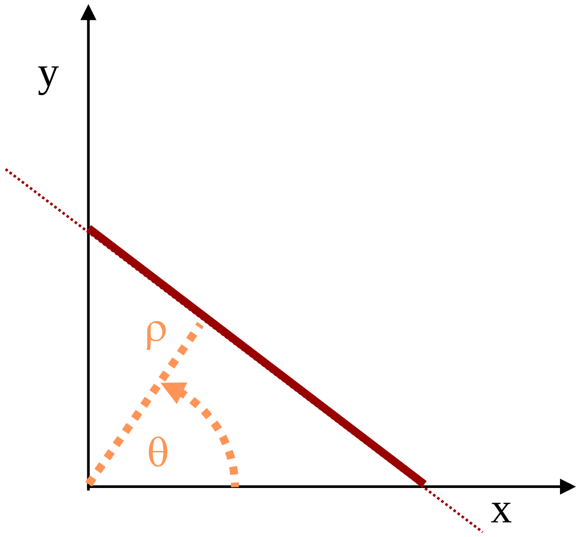
HT: searching straight lines

- In the classic equation the parameters are not limited:

$$-\infty < m, q < +\infty$$

- For this reason Hough adopted a different straight line representation introducing the a PS (ρ, θ):

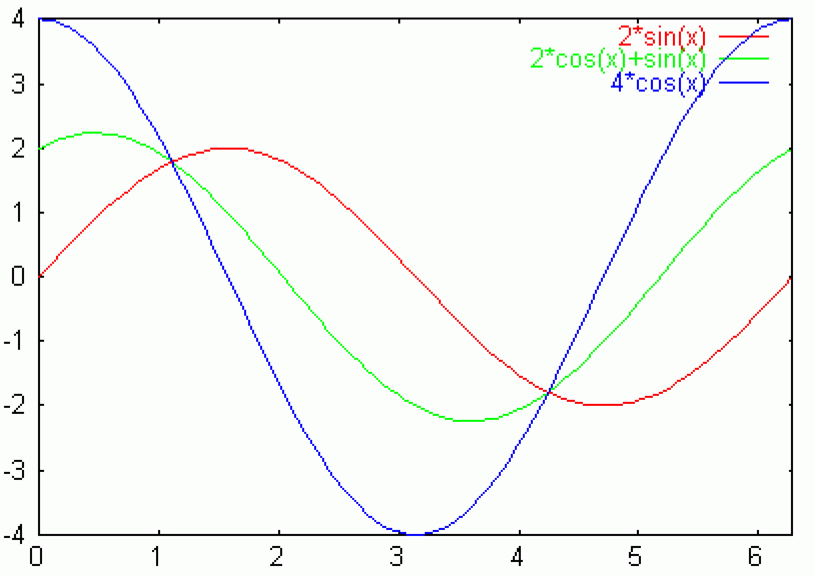
$$\rho = x \cos(\theta) + y \sin(\theta).$$



HT: searching straight lines

- In this case the PS is limited to:

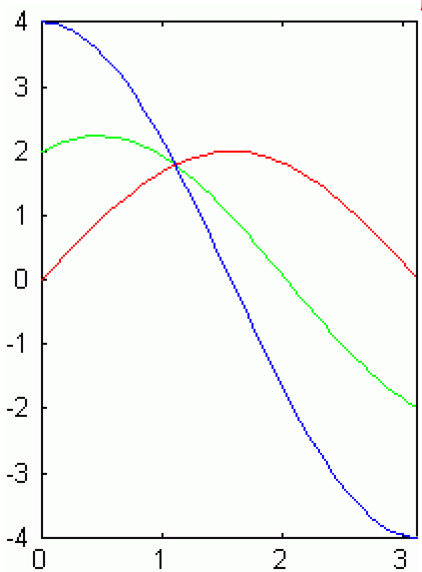
$$0 < \rho < L\sqrt{2}; -\pi \leq \theta \leq \pi$$



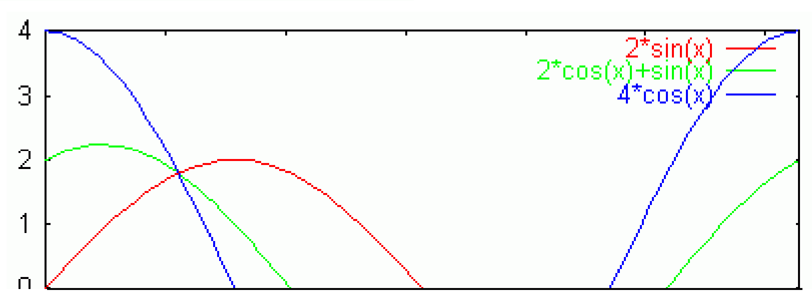
HT: searching straight lines

- In this case the PS is limited to:

$0 < \rho < L\sqrt{2}; -\pi \leq \theta \leq \pi$

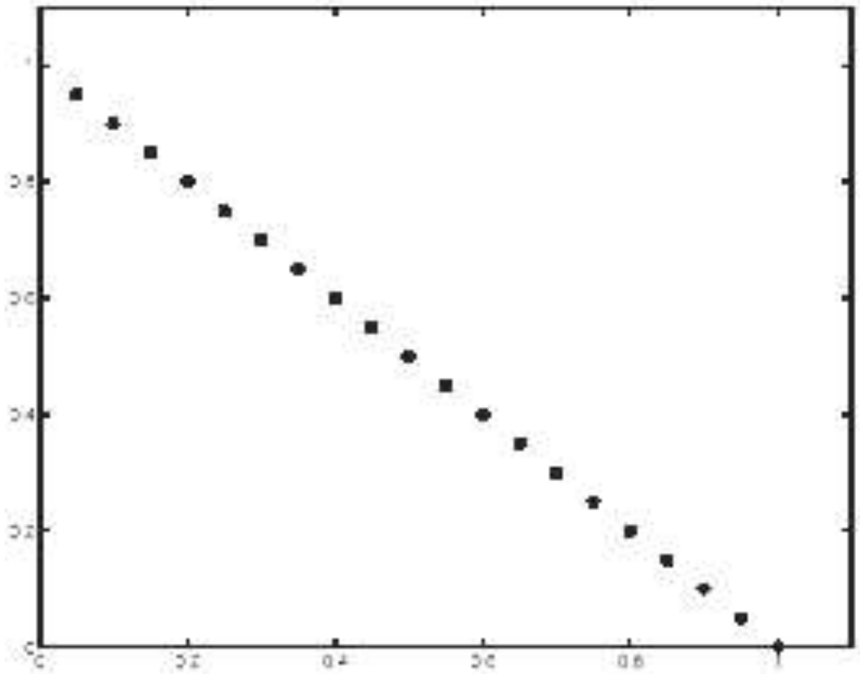


$-R\sqrt{2} \leq \rho \leq R\sqrt{2};\ 0 \leq \theta < \pi$

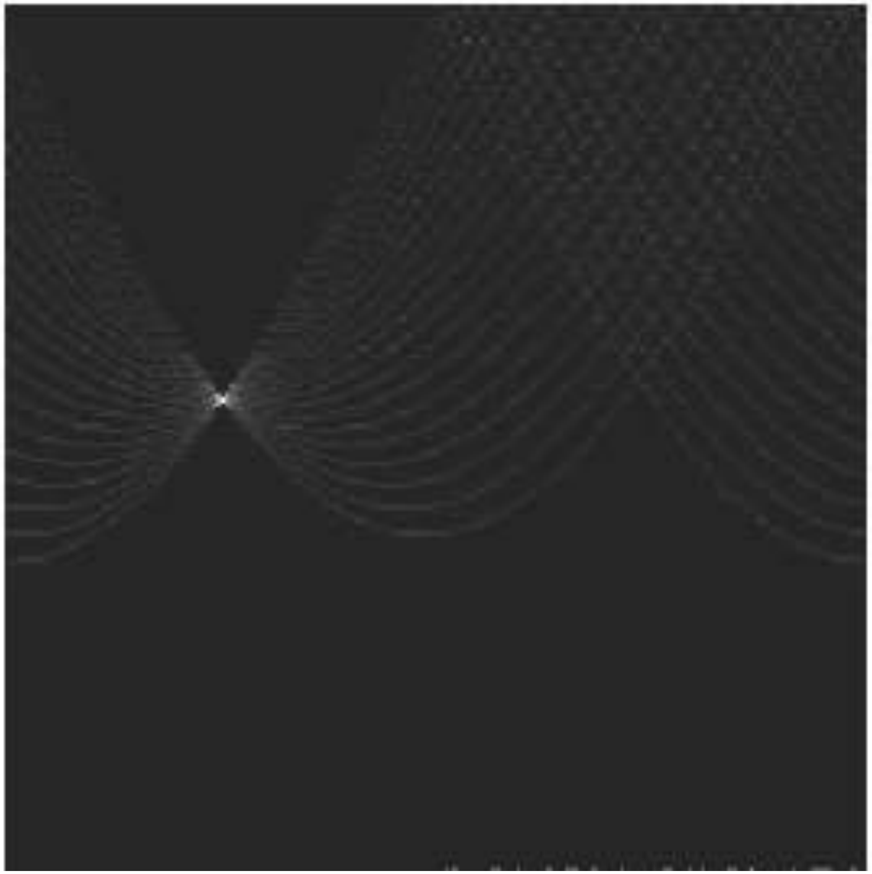


$) \leq \rho \leq R\sqrt{2};\ 0 \leq \theta < 2\pi$

Hough transform - experiments



features

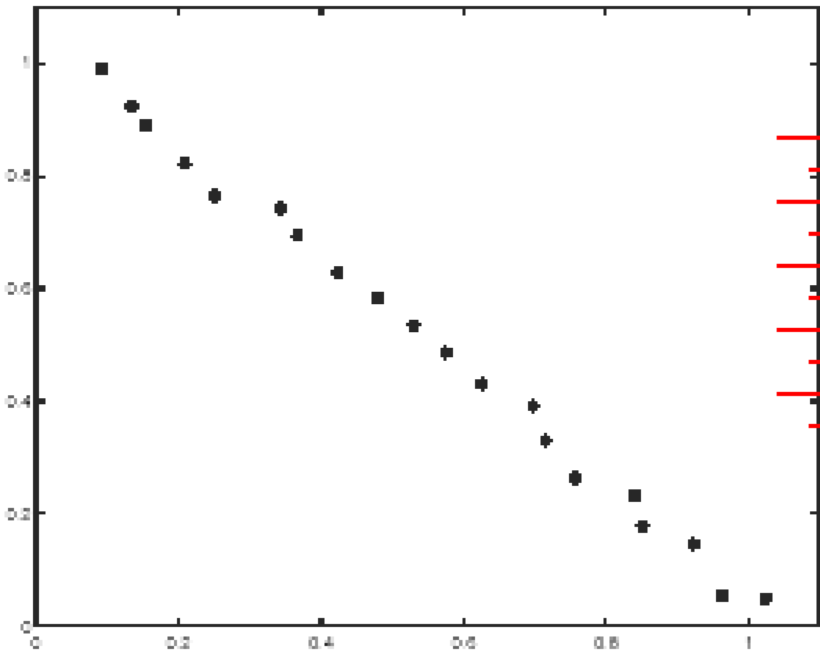


votes

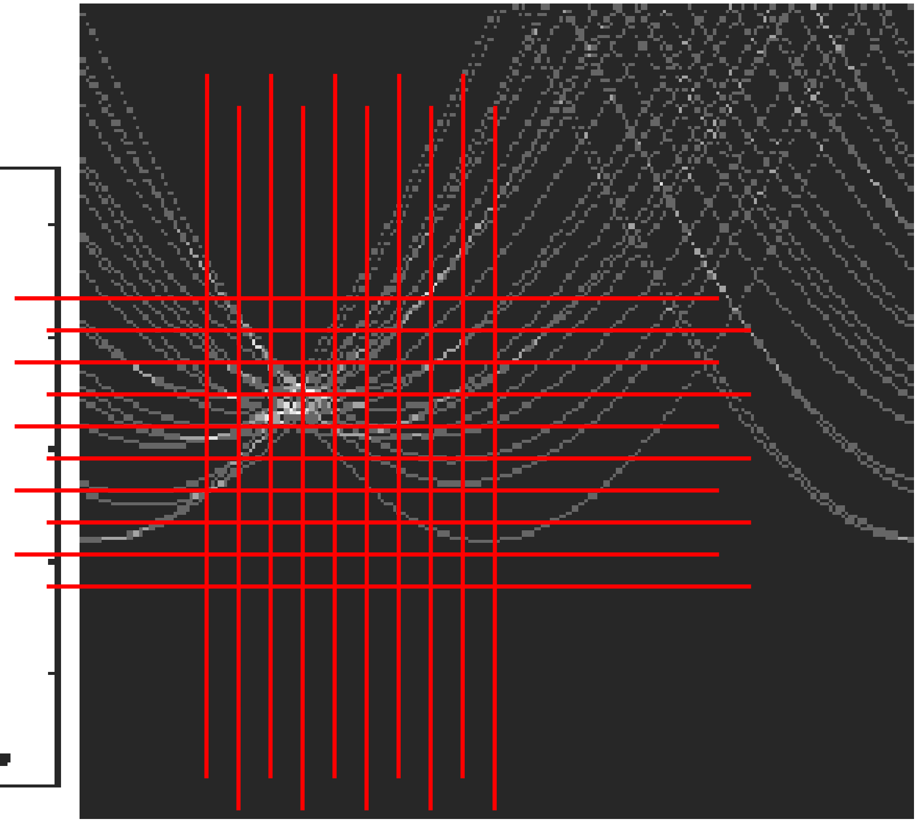
Slide from S. Savarese

Hough transform - experiments

Noisy data



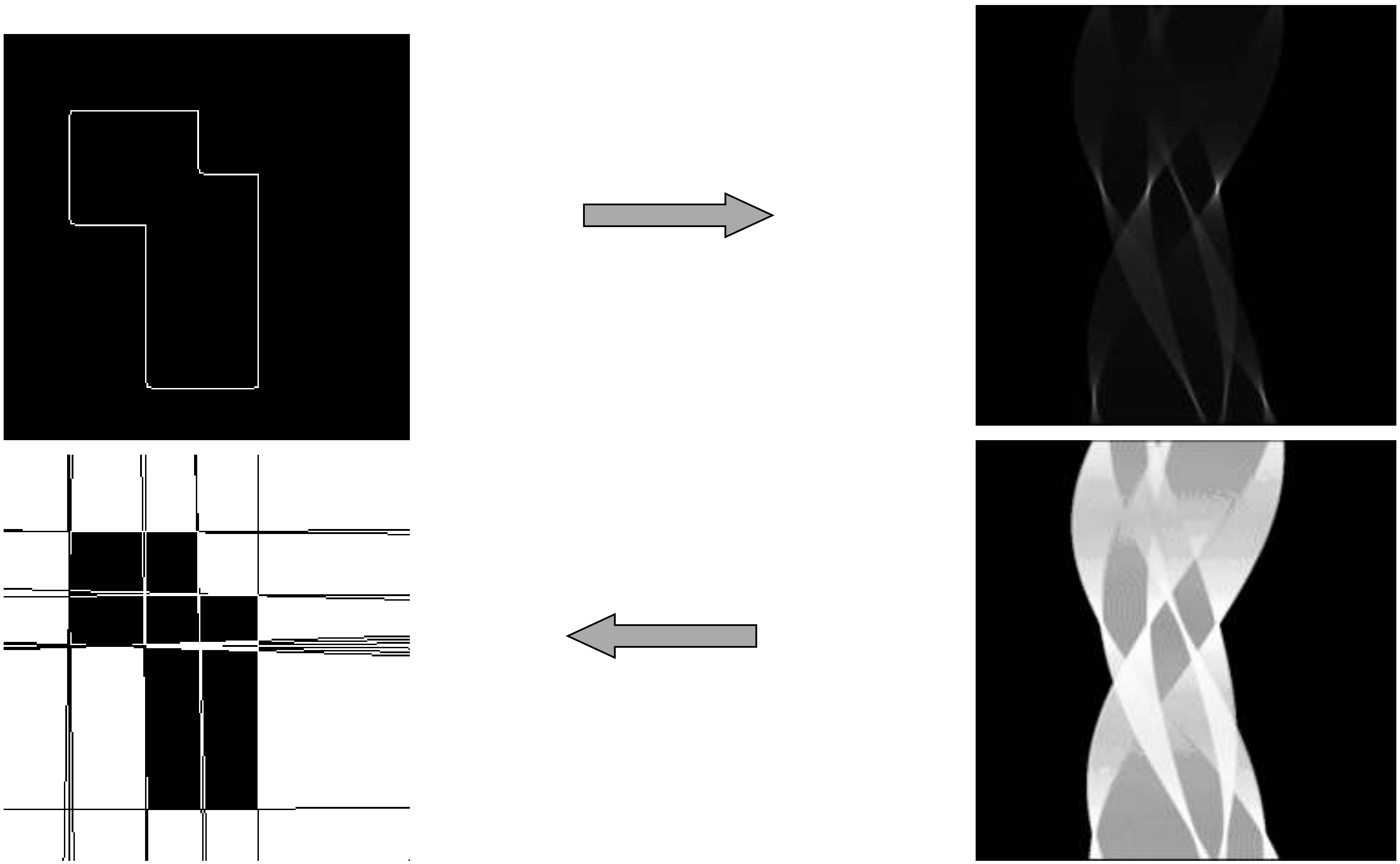
features



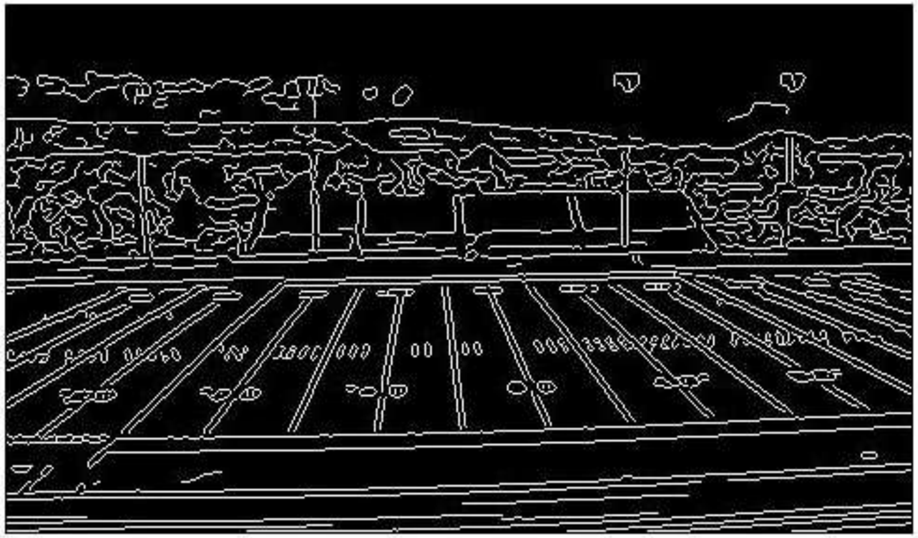
votes

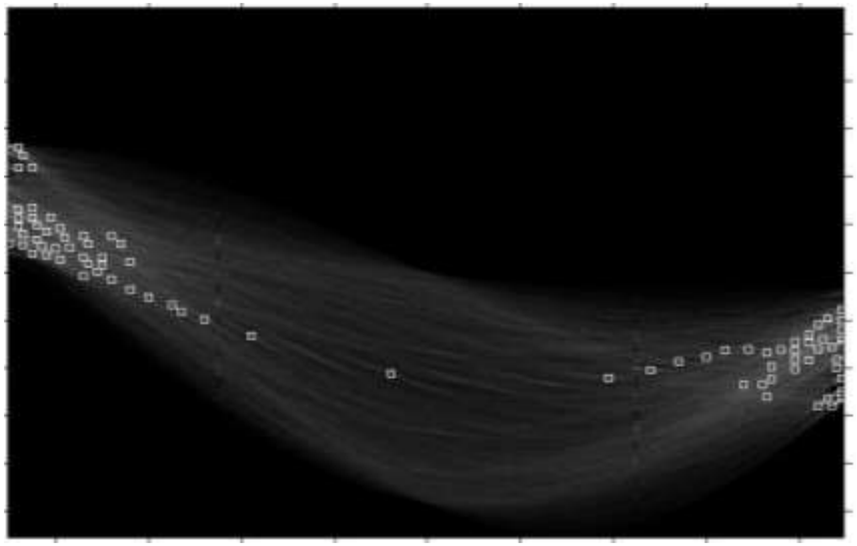
Need to adjust grid size or smooth

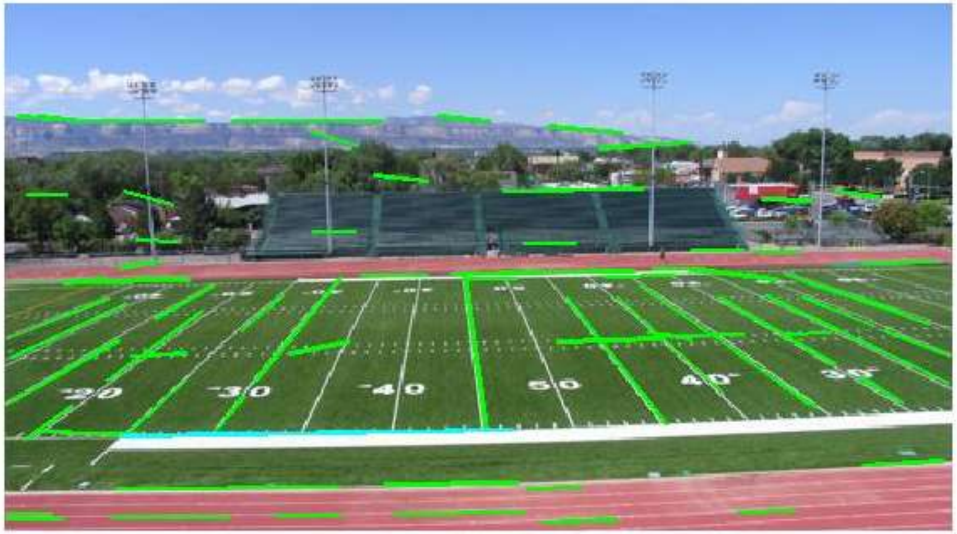
Line Detection by Hough Transform



Showing longest segments found

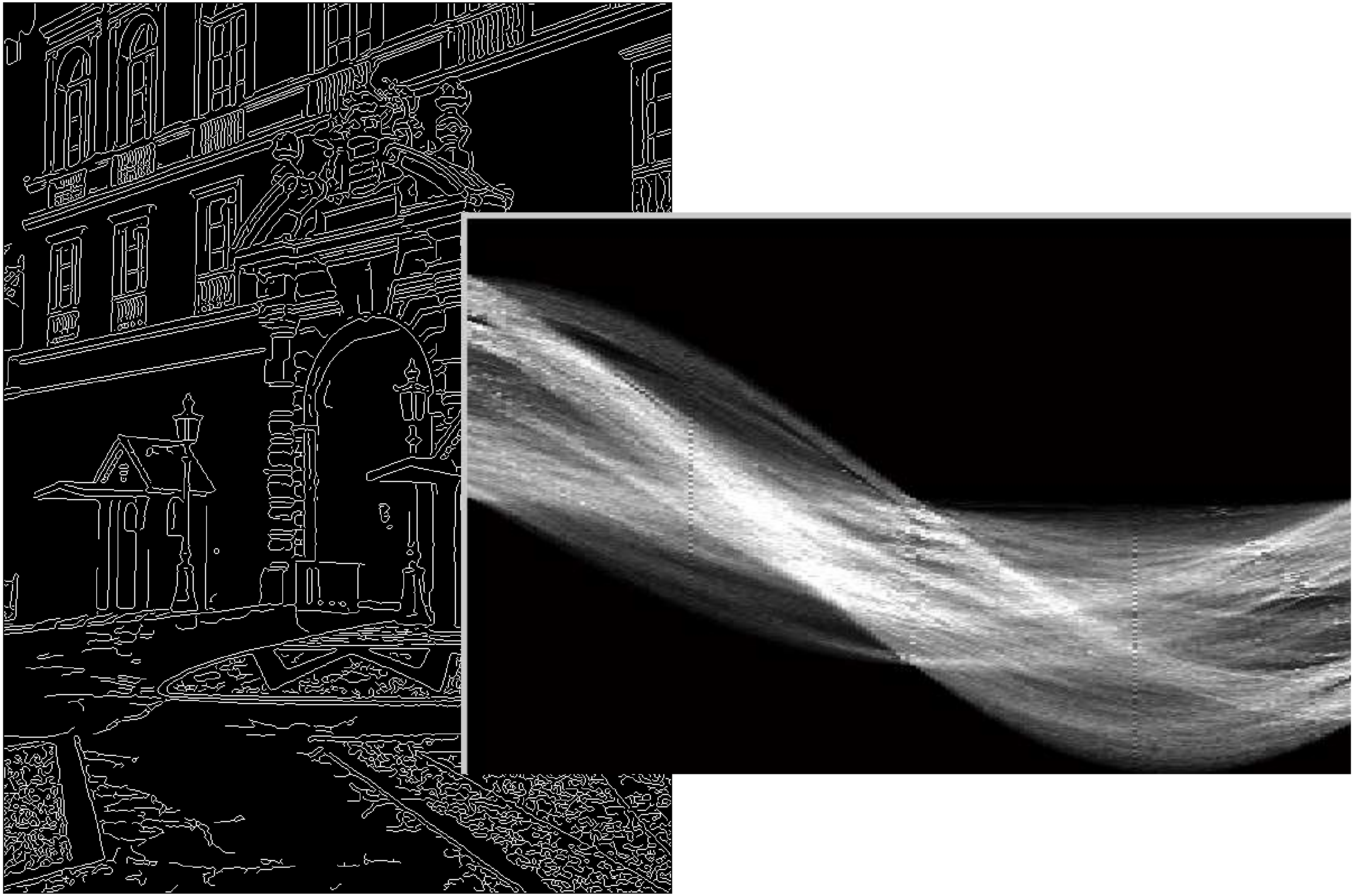






1. Canny edge detection


2. Edge points → Hough votes



3. Hough votes → Edges

Find peaks and post-process

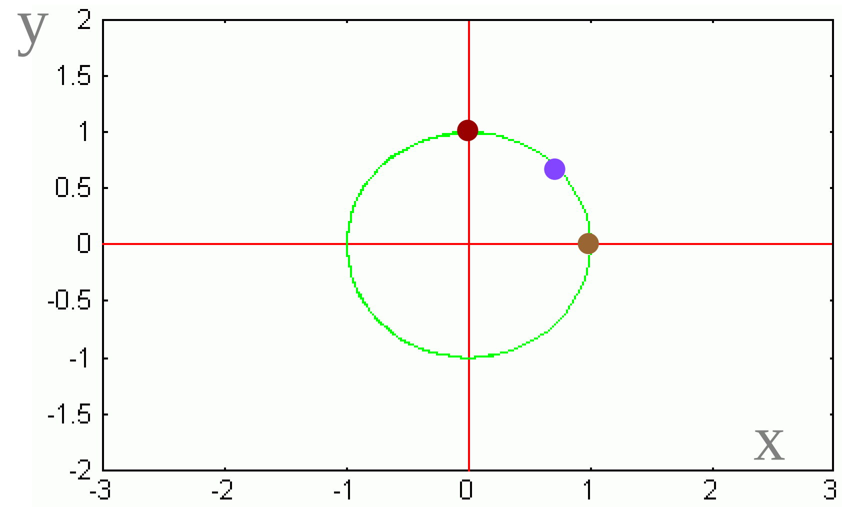
An example

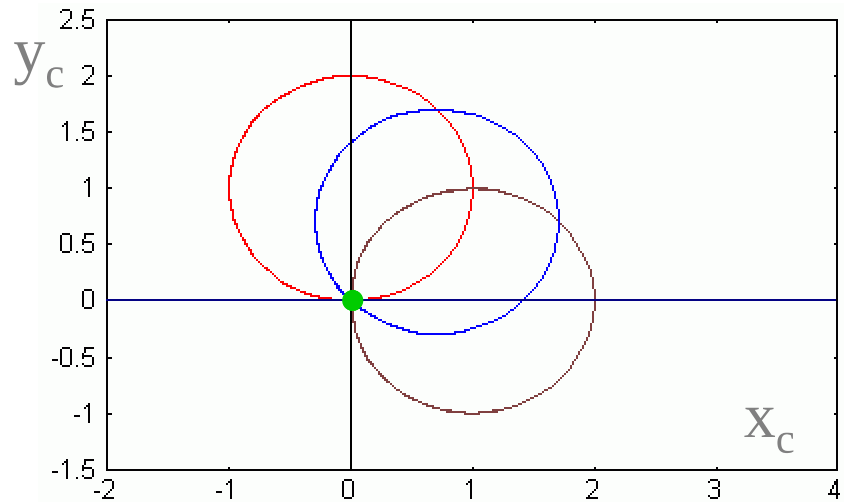
HT: searching analitycal curves

- From the very first Hough transform version, many extensions were developed along the years. It ranges from more complex analytical curves (with a higher number of parameters), e.g. circles:

$$(y-y_c)^2+(x-x_c)^2=r^2$$

- First case: search circles with a given radius.
  - we have a 2D PS which represents the circle center coordinates (x_c,y_c)
  - the mapping rule (locus of compatible points) is also a circle with the given radius.
    Note that it is not always true that searched curves and mapping rule are equal

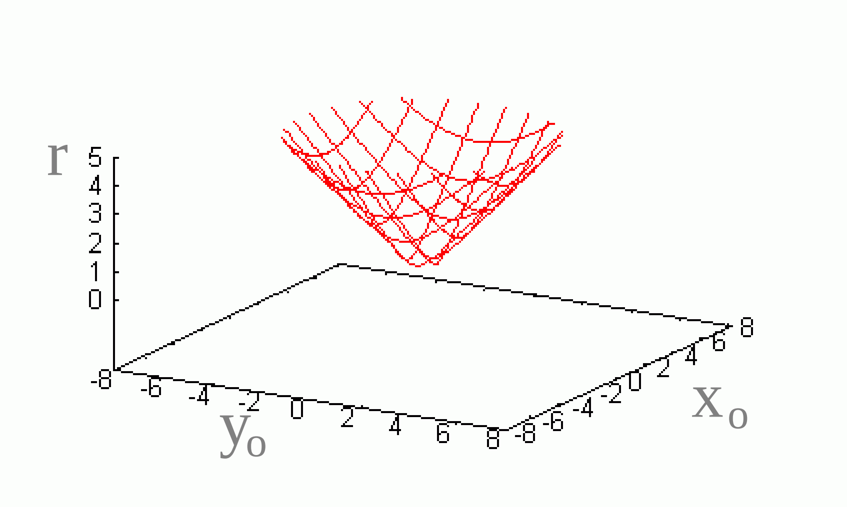




HT: searching for circles

- If the radious is unknown the PS is 3D:

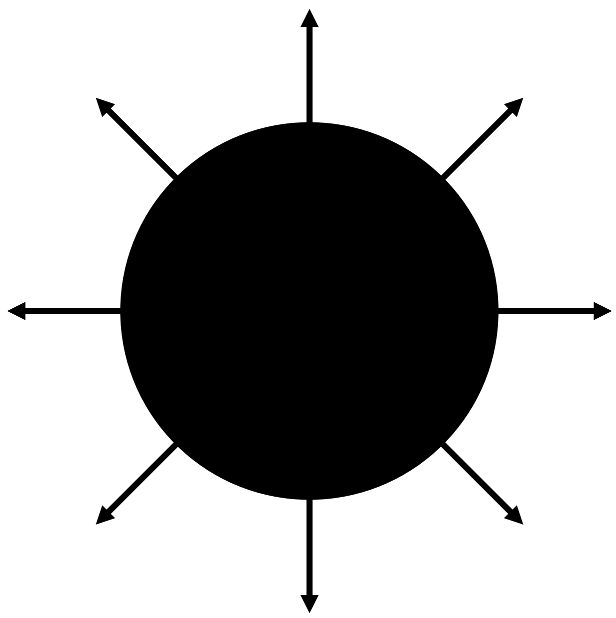
  $f((x,y),(x_c,y_c),r) = (y-y_c)^2+(x-x_c)^2-r^2=0$
- The mapping rule is a cone.

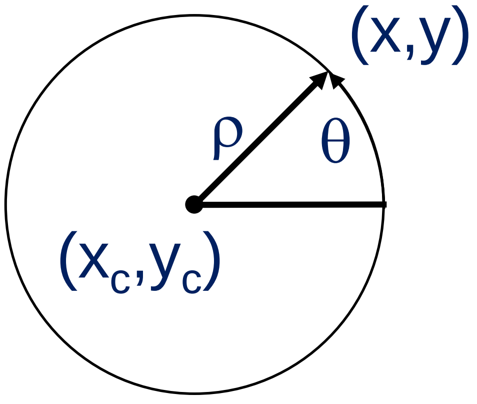


Using orientation

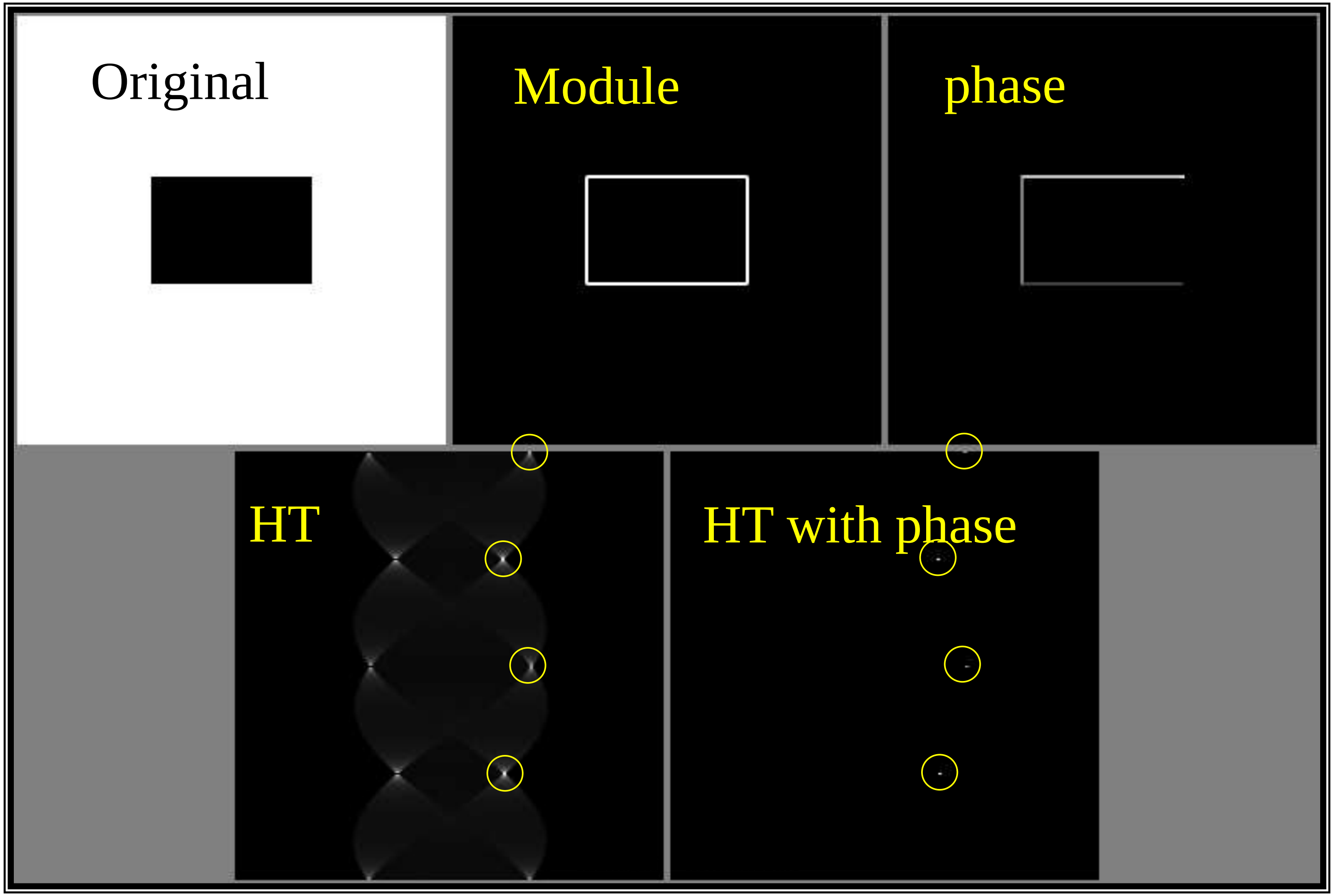
- An edge detector may also provide the contour orientation; in this case the possible detection is more effective with less spurious artifacts.
  - The number of independent parameters decrease by one
- Examples:
  - Straight lines: $\rho = x \cos(\theta) + y \sin(\theta)$
    only the point (ρ,θ) in the PS is voted
  - Circle hough: $x_c = x - \rho \cos(\theta)$; $y_c = y - \rho \sin(\theta)$
    only the point (x_c, y_c) in the PS is voted

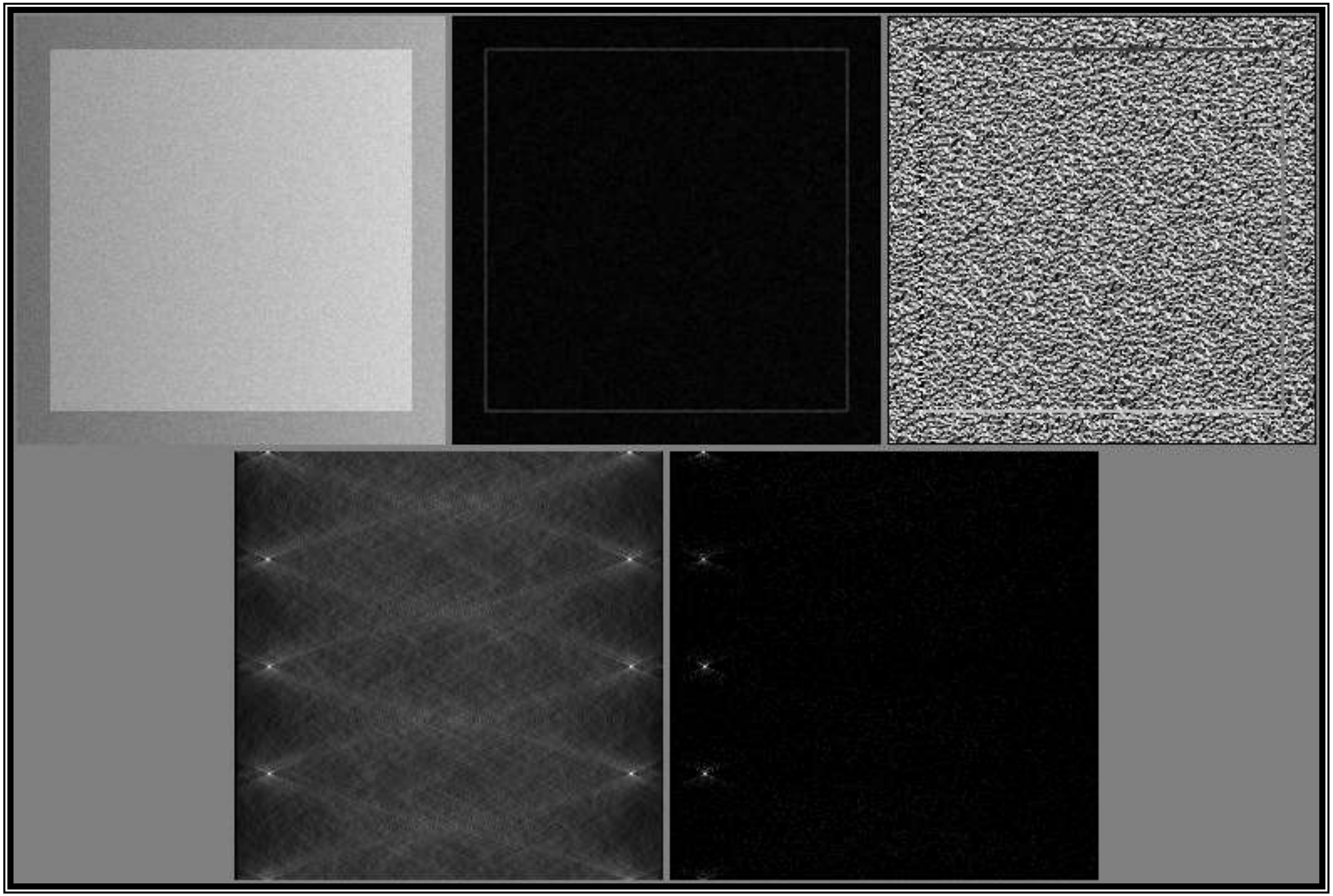
Example circle hough

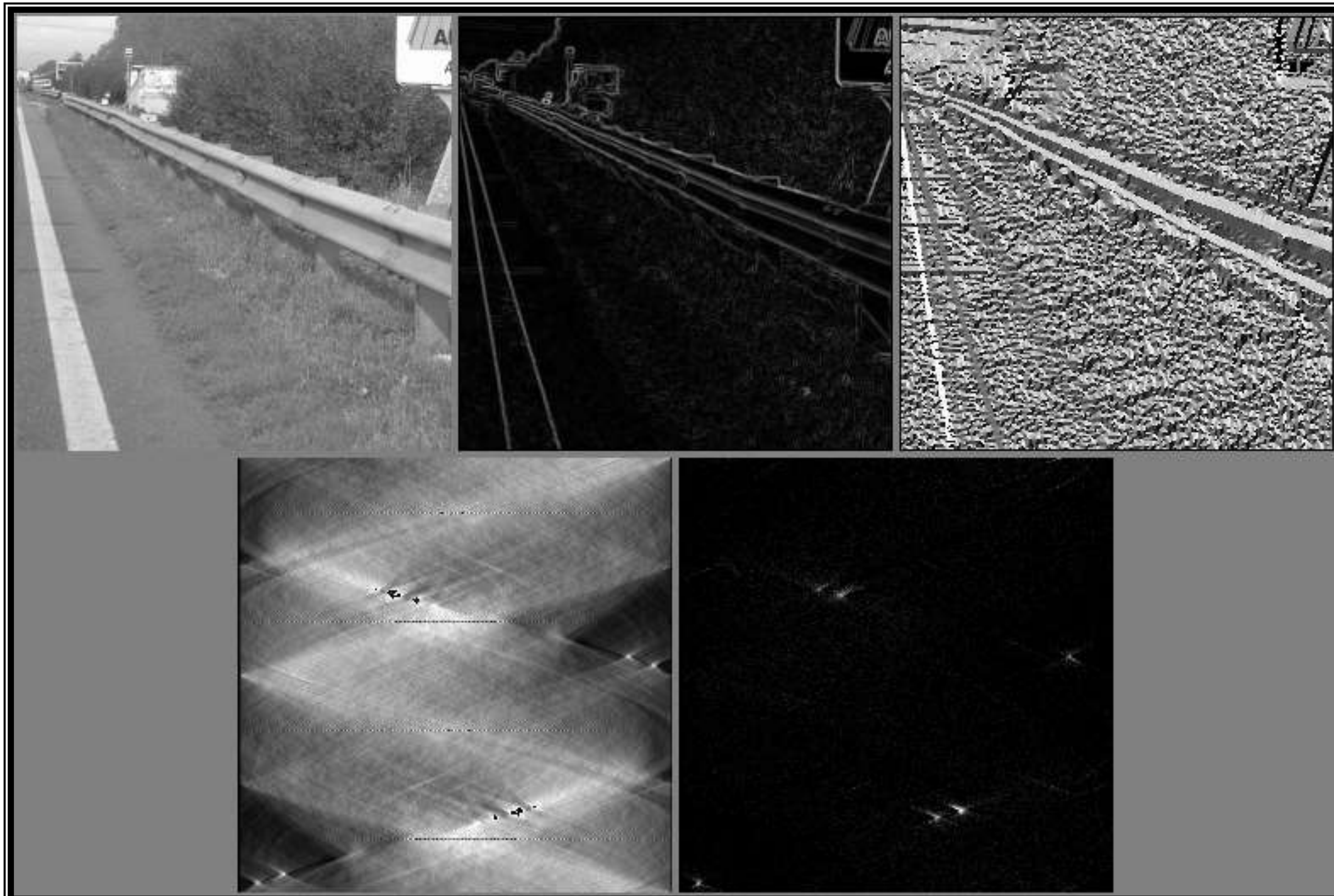




$$x_c = x - \rho \cos(\theta)$$
$$y_c = y - \rho \sin(\theta)$$


Original
Module
phase
HT
HT with phase





Implementation of the HT

- The original approach of the HT is based on these elements:
  - an **enriched edge detector** to find contour pixels and some local properties **as the gradient angle or local curvature** (concavity and convexity);
  - an array (in a **parameter space**) working as an accumulator of the contributions. **Each element of the parameter space represents a possible instance of the searched object**;
  - a **mapping rule** which defines the contributions of the detected instance on the accumulator array.
    - ✓ The simplest solution is to **increment all the elements**, corresponding to the pattern, **compatible with the detected instance**.
    - ✓ A weighted contribution can be introduced **on the basis** of both **the estimated precision** (e.g. the further the location the lower the contribution because of the edge detection orientation bias) **and/or of the saliency of the detected instance**;
  - **a discriminant criterion for the evaluation of the resulting final contribution in the parameter space**. Knowing **the expected maximum contribution**, the common solution is by a threshold (local maxima over the threshold identify the presence of the pattern), so taking care of possible occlusions, noise, etc.

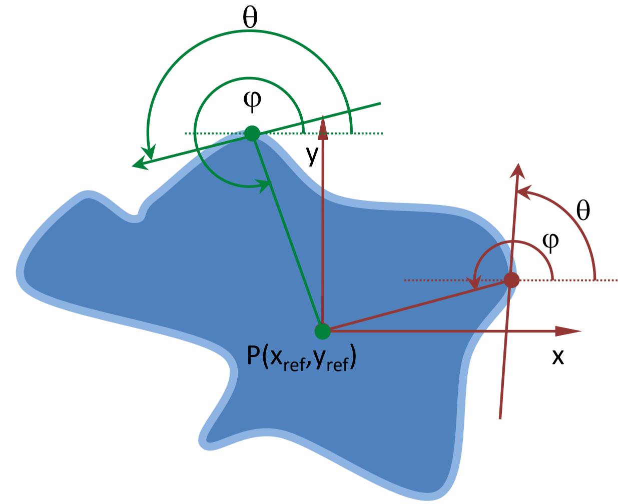
The Generalized Hough Transform

- With the ***Generalized Hough Transform* (GHT)**, under the assumption of rigid motion, any pattern can be represented and recognized.
- Let us first consider the case of a pattern given as silhouette at fixed scale
- Let us select a reference point $P_{ref}(x_{ref}, y_{ref})$, not essentially the center of gravity (moments of firdt order), even if its centrality is often advantageous
- Each boundary point $P(x_o, y_o)$ can be referred to P_{ref} as:

$$\rho = \sqrt{\left(x_{ref} - \overset{o}{x}\right)^2 + \left(y_{ref} - \overset{o}{y}\right)^2}$$

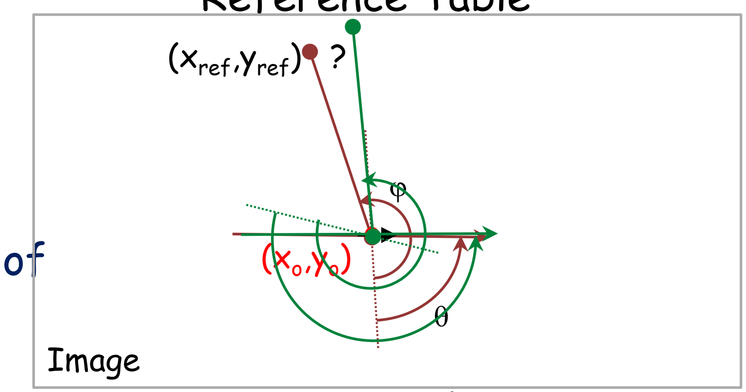
$$x_{ref} = \overset{o}{x} + \rho\cos(\varphi - \theta) \quad y_{ref} = \overset{o}{y} + \rho\sin(\varphi - \theta)$$

- The **mapping rule** that corresponds to the set of object contour points, can be described by a **Reference Table (RT)** with the illustrated geometry



....	....	....	....
P(x,y)	*ρ*	*φ - θ*	*other peculiarities*
....	....	....	....

Reference Table

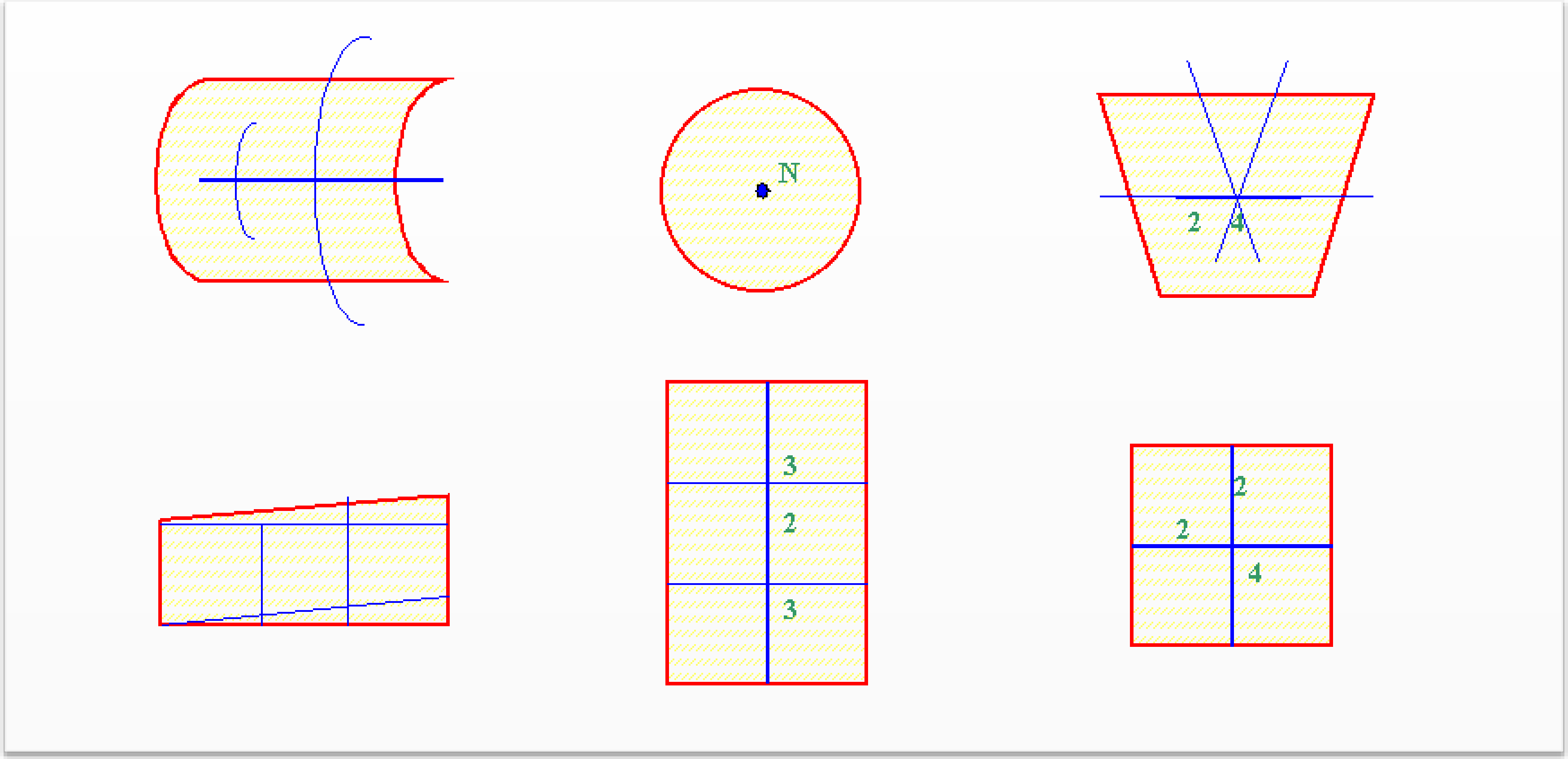


Mapping rule

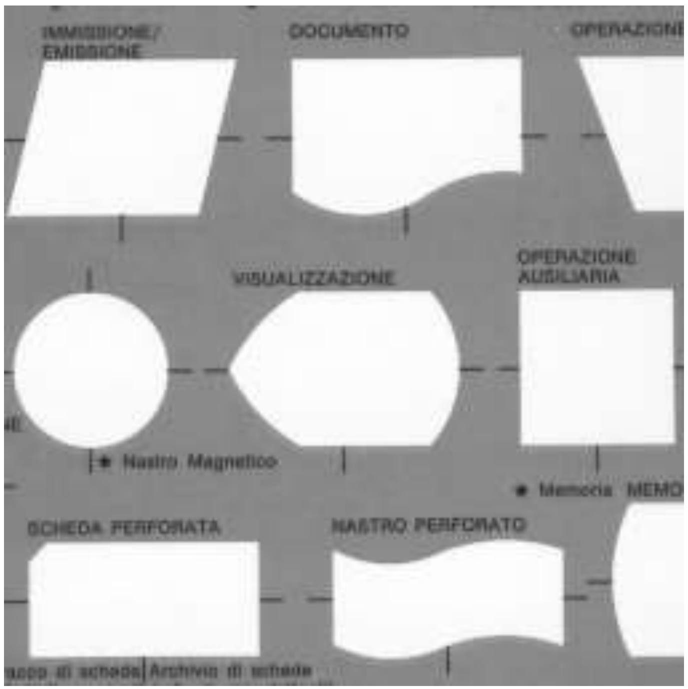
The Generalized Hough Transform

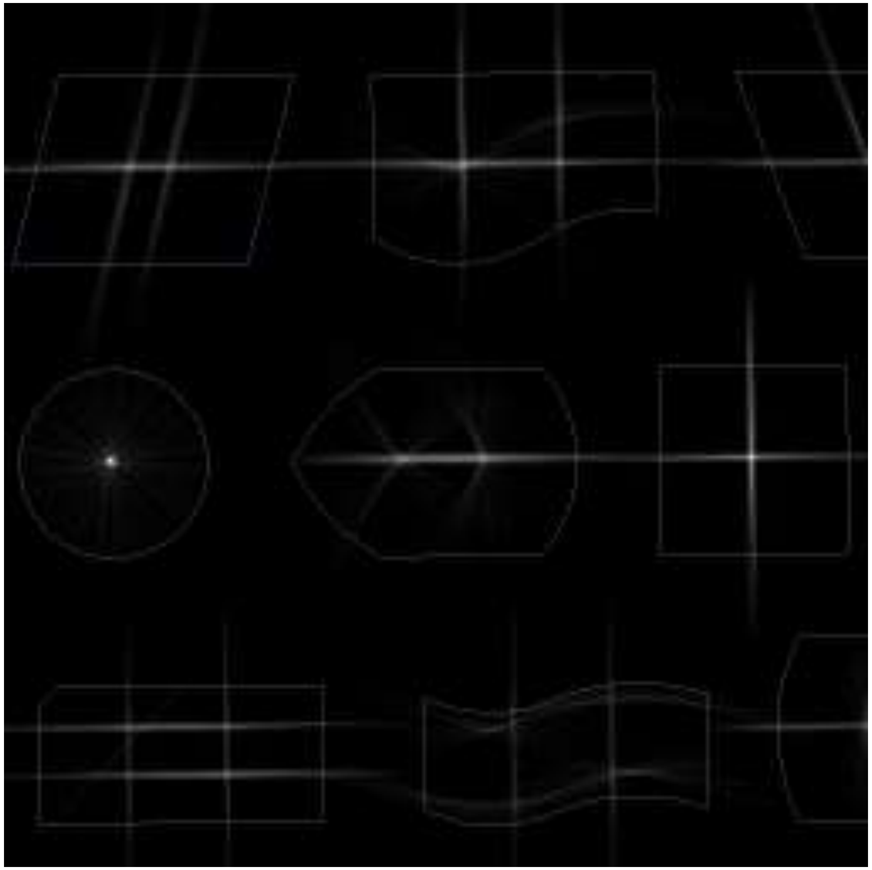
- For a given point P(x,y) obtained by an edge detector on the image under analysis, the set of compatible points of the PS (which here represent the positions of the P_{ref} compatible with the contour crossing P and the **PS coincide with the image space**) are given by the equation above that represents the mapping rule.
- In the case of fixed size object, for each contour point detected on the image, **the number of contributions onto PS is N, the cardinality of the RT**. If all the contour points of a searched pattern, are effectively present in the image are detected properly, **a peak of value N will appear in the P_{ref} position of PS** (corresponding to the model roto-traslation)
- Instead, considering the case of unknown scale factor **s**, to the image describing the position of the P_{ref} in the image, **an extra dimension must be introduced in the PS for the parameter s**. It become a 3D PS (replicating the image for each value of **s**) and in the above equation ρ must be multiplied by the correspondent scale factor **s**.
- In a similar way, if we want to detect directly the object orientation, for a 2D object, all the process must be realized in a 4D PS: $(x_{ref}, y_{ref}, s, \varphi)$. **As we will see later a more convenient solution is to choose a couple of P_{ref}.**

Example



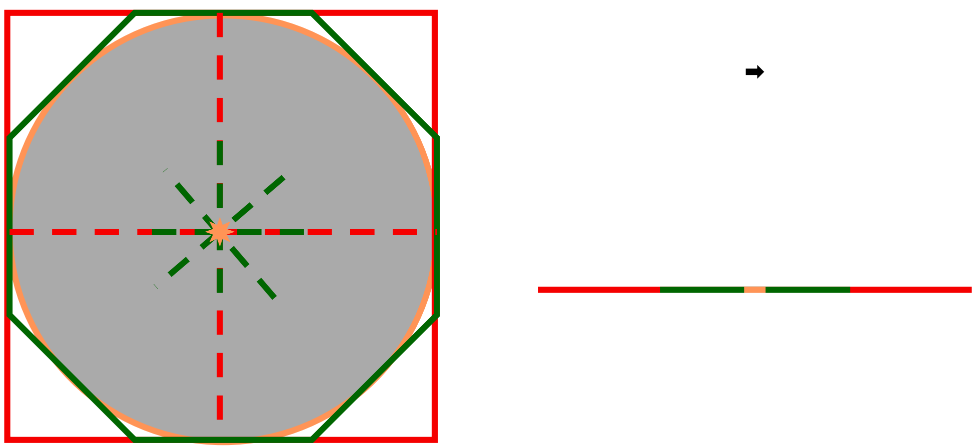
Example: looking for a square



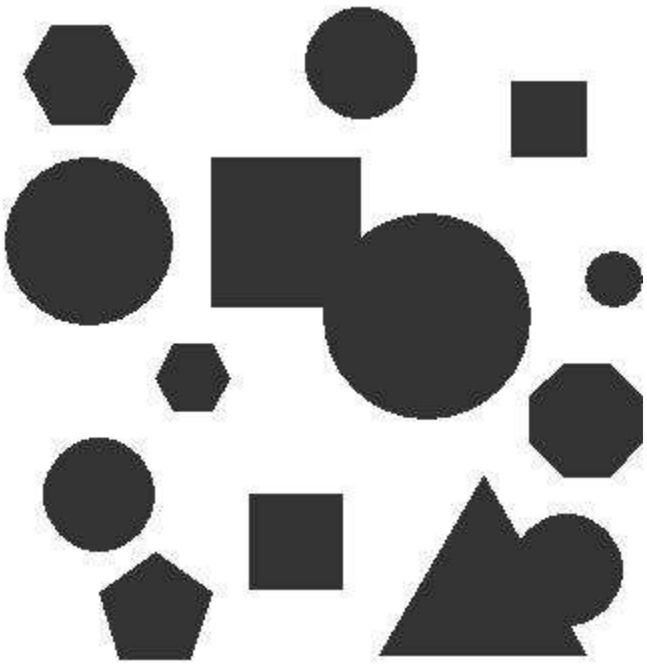


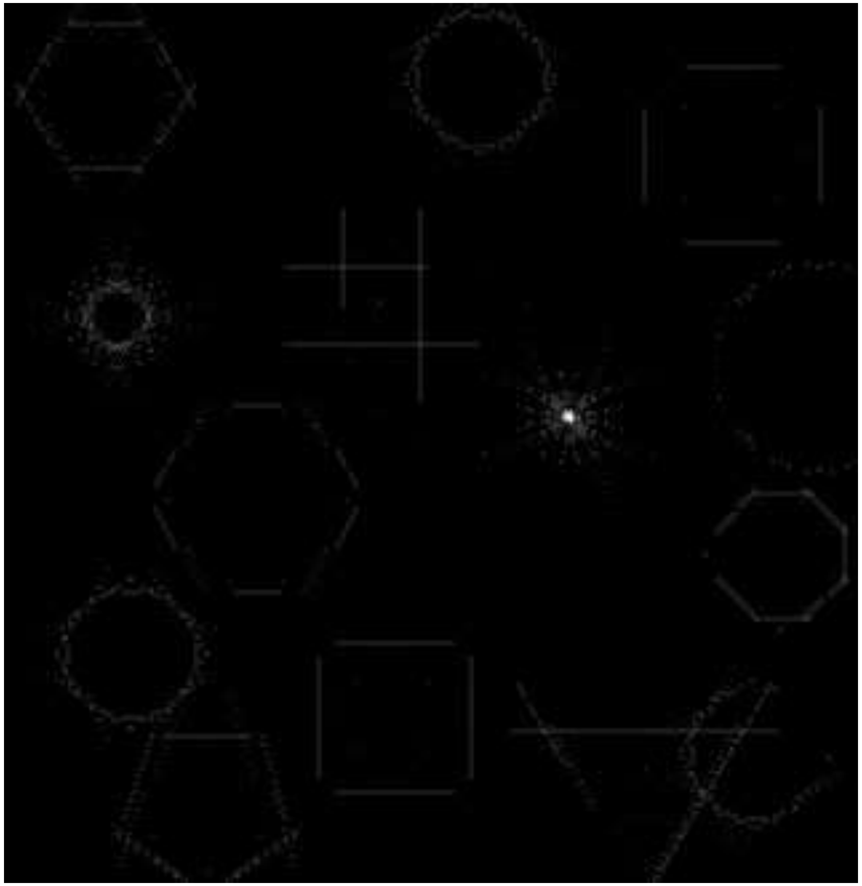
Regular polygons

- Regular polygons have a mapping rule consisting in a side distant from the center as the apothem
- The mapping rule of a regular polygon of order n contains all the mapping rules of regular polygons of higher order having the same apothem
- Looking for a regular polygon of order n, it will gather a number of votes V equal to its perimeter: $V = n\ L_n$
- If another regular polygon of order m with the same apothem is present it will gather a number of votes V equal to: $V = m\ L_n$ if $m \leq n$, $V = m\ L_m$ if $m > n$ but note that in this case $nL_n > m\ L_m$

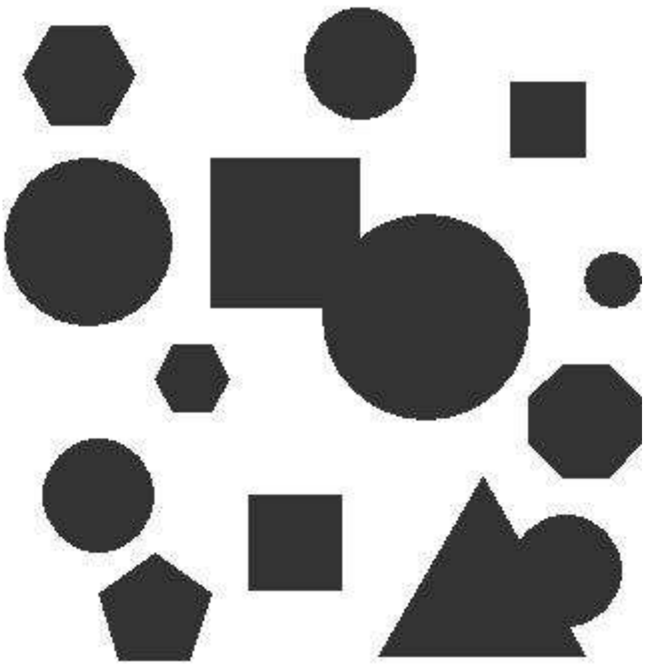


Example: looking for a circle

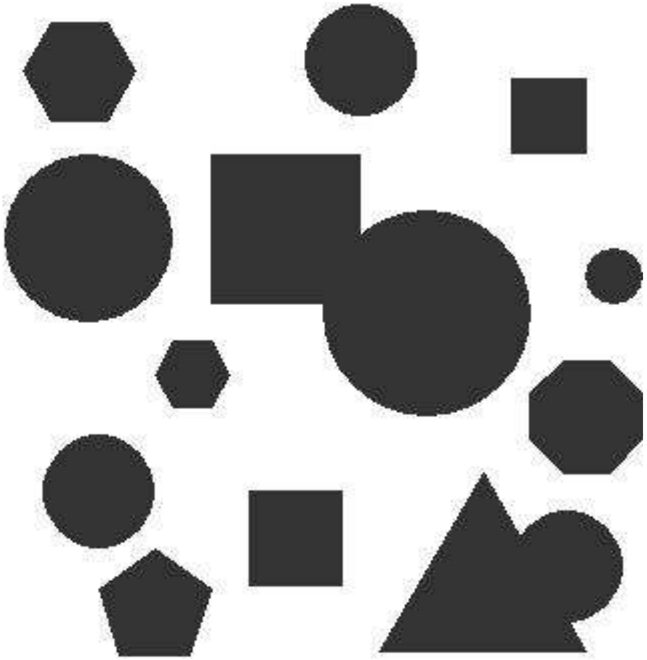




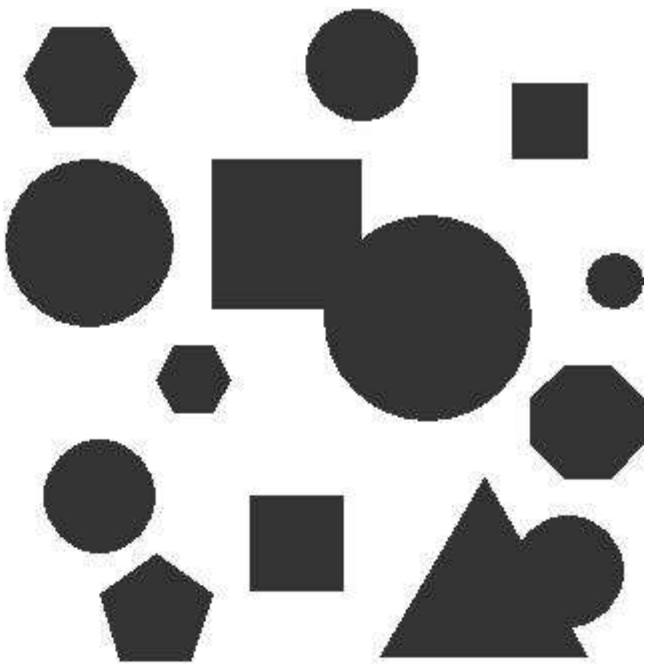
Example: looking for a octagon



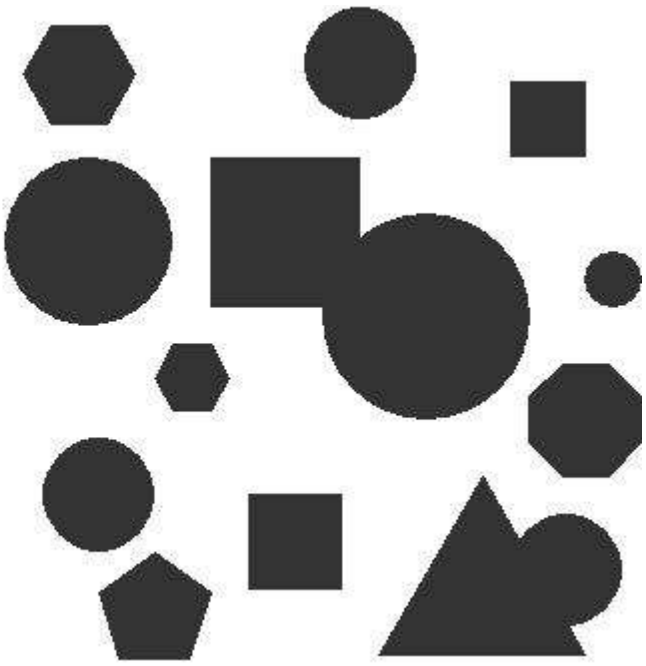
Example: looking for a hexagon



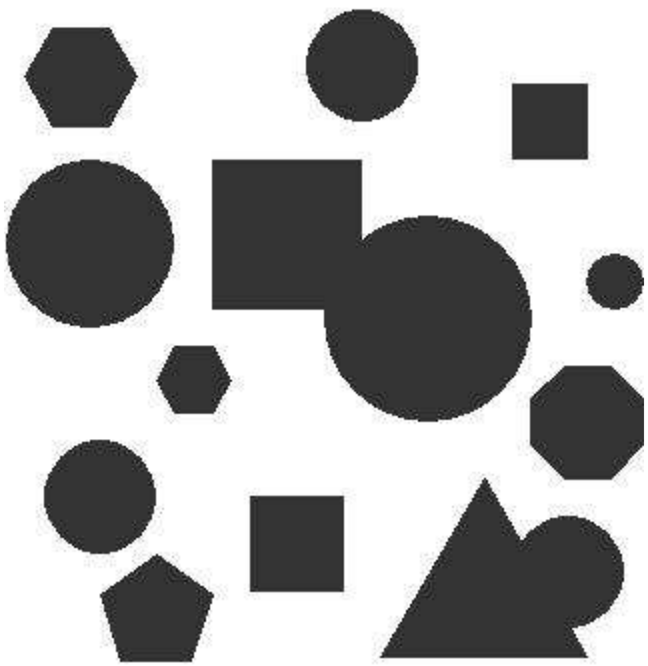
Example: looking for a pentagon

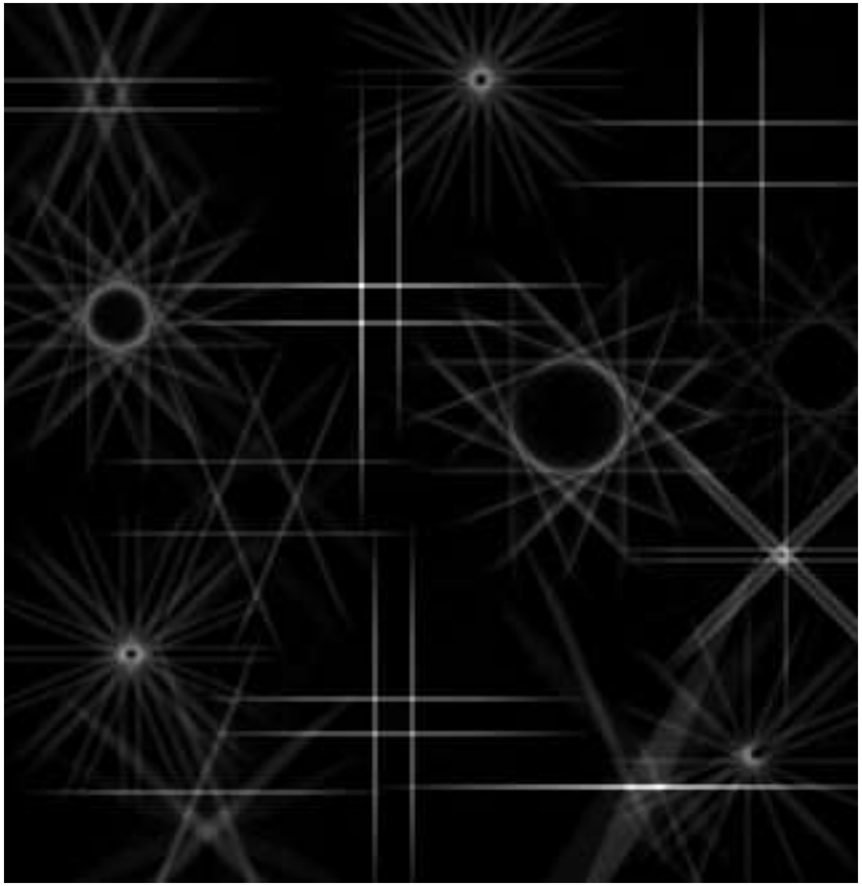
 

Example: looking for a square

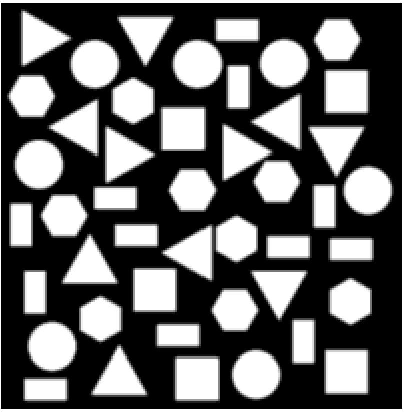


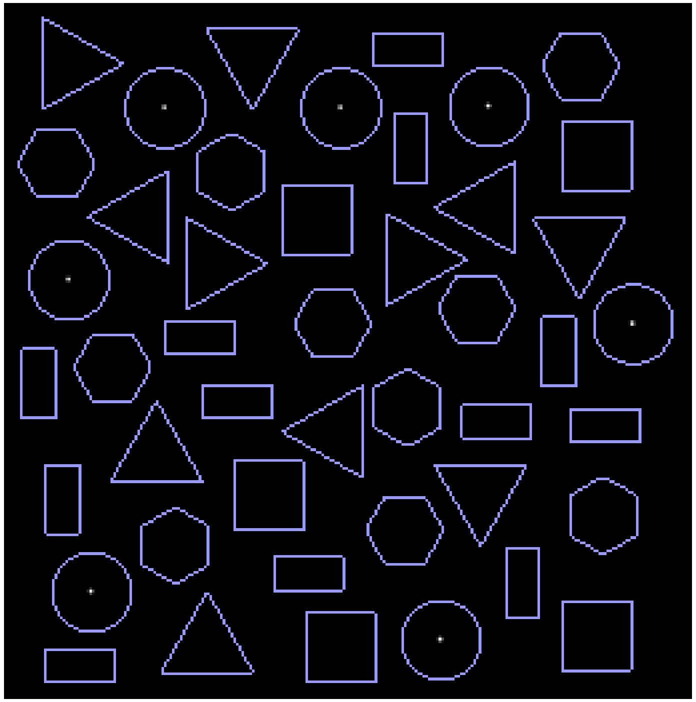
Example: looking for a triangle

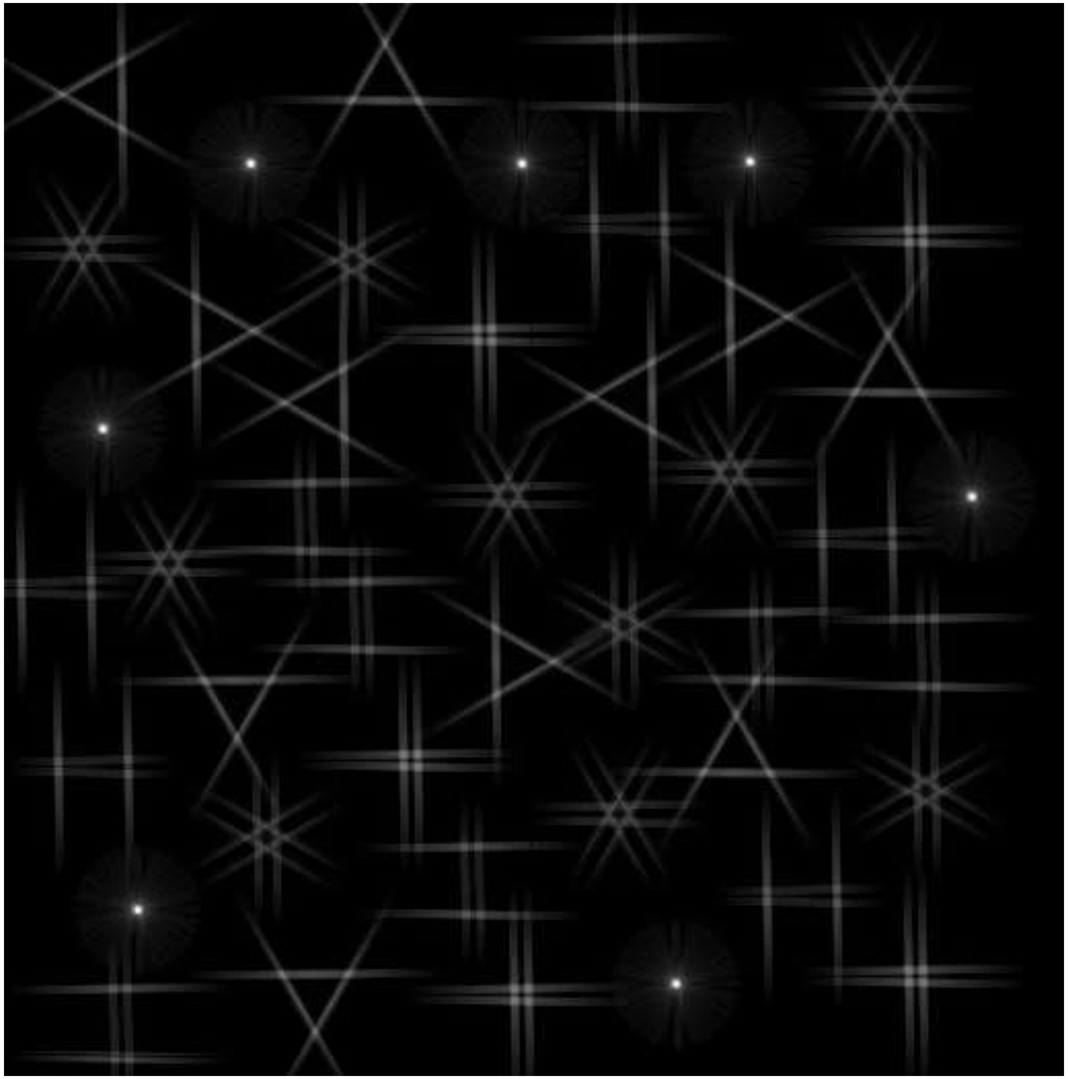




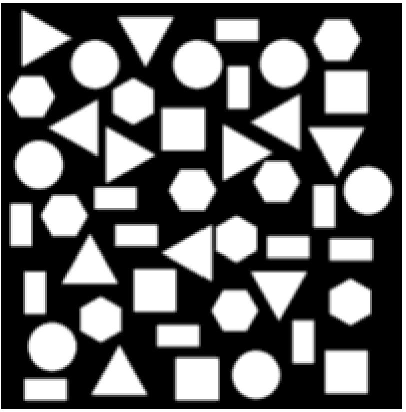
looking for a circle

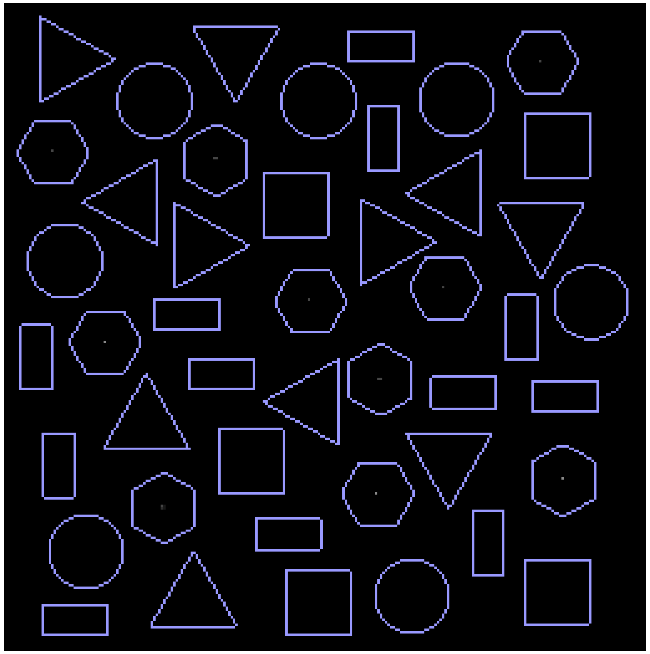


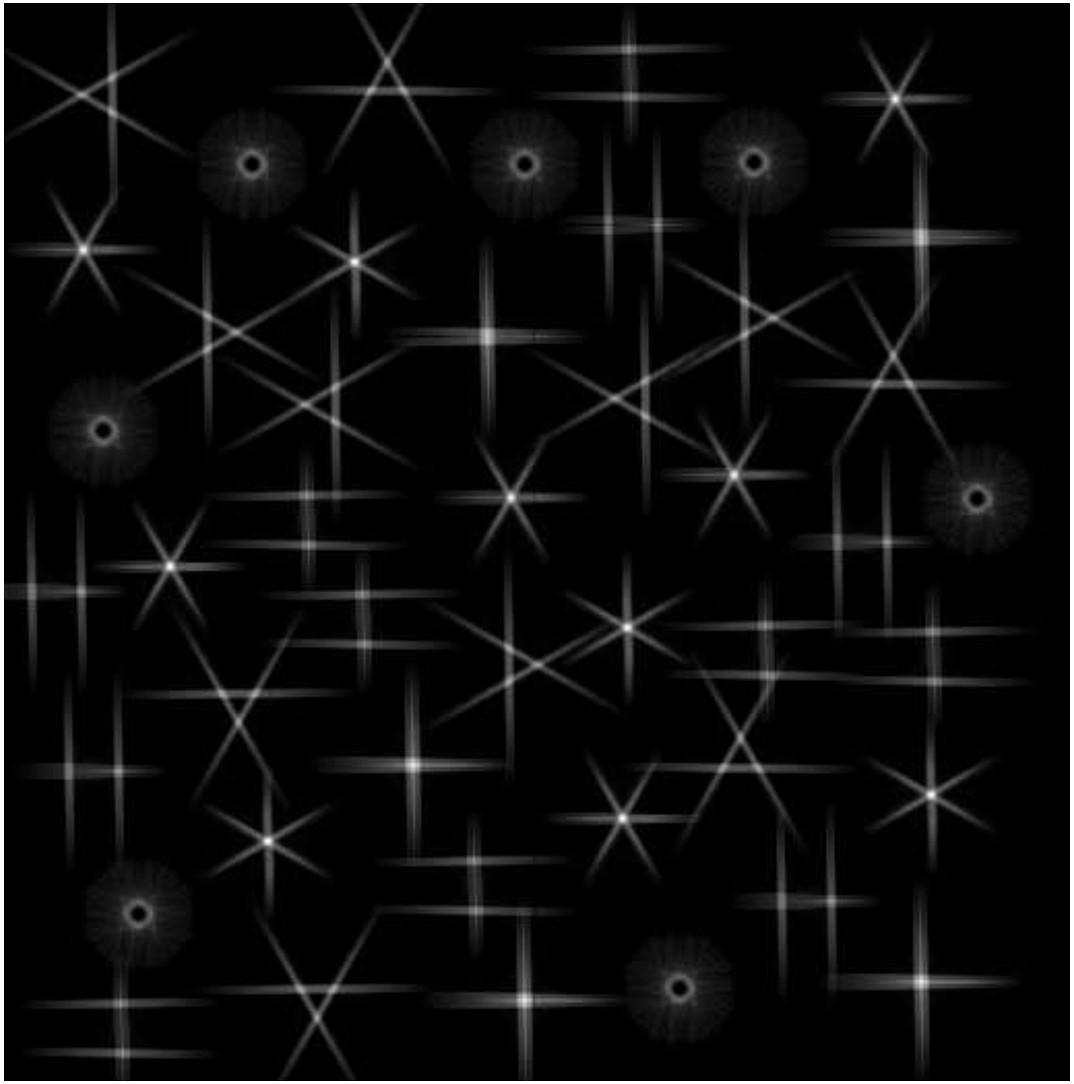




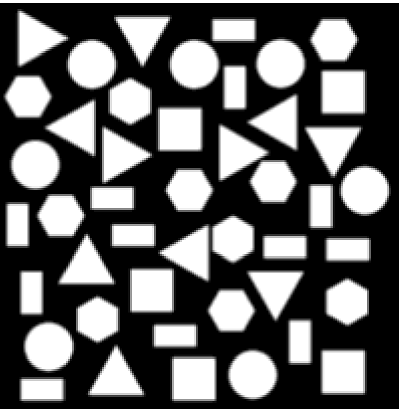
looking for a hexagon

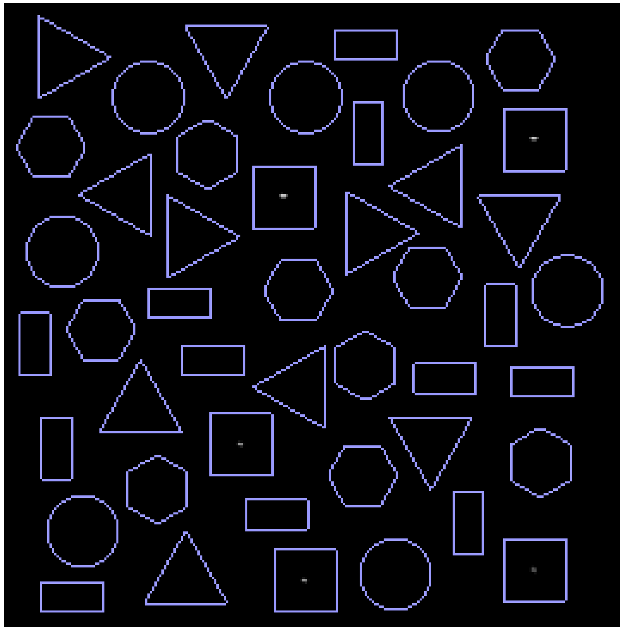


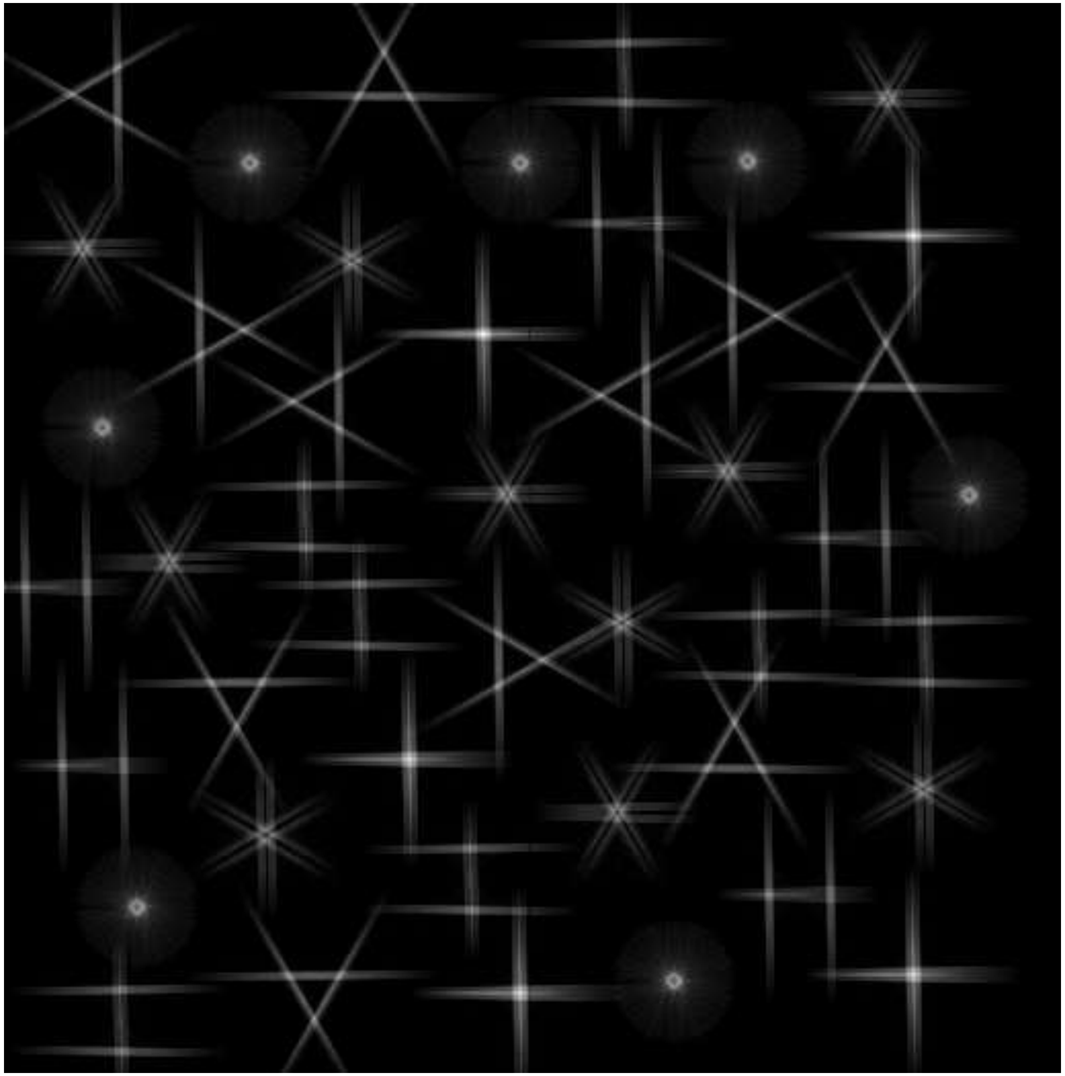


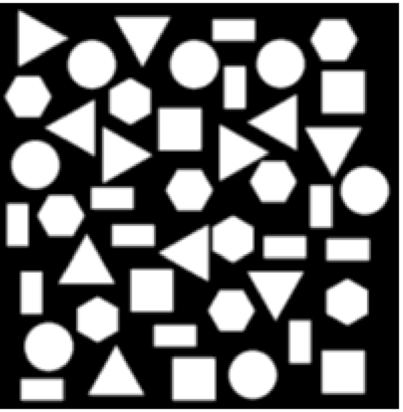


looking for a square

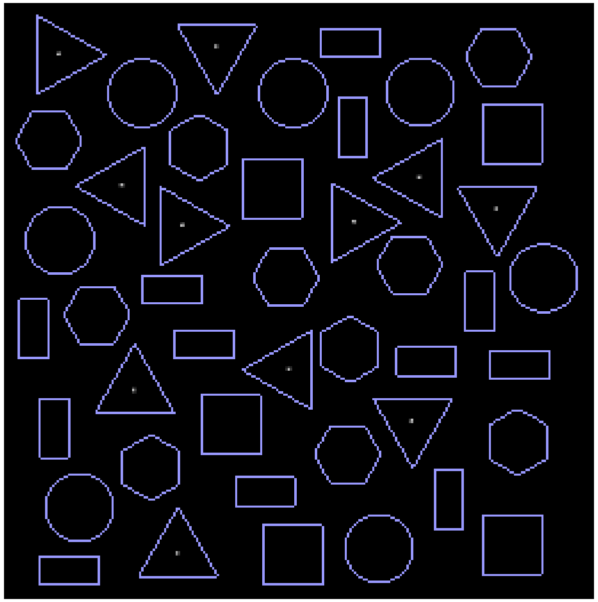


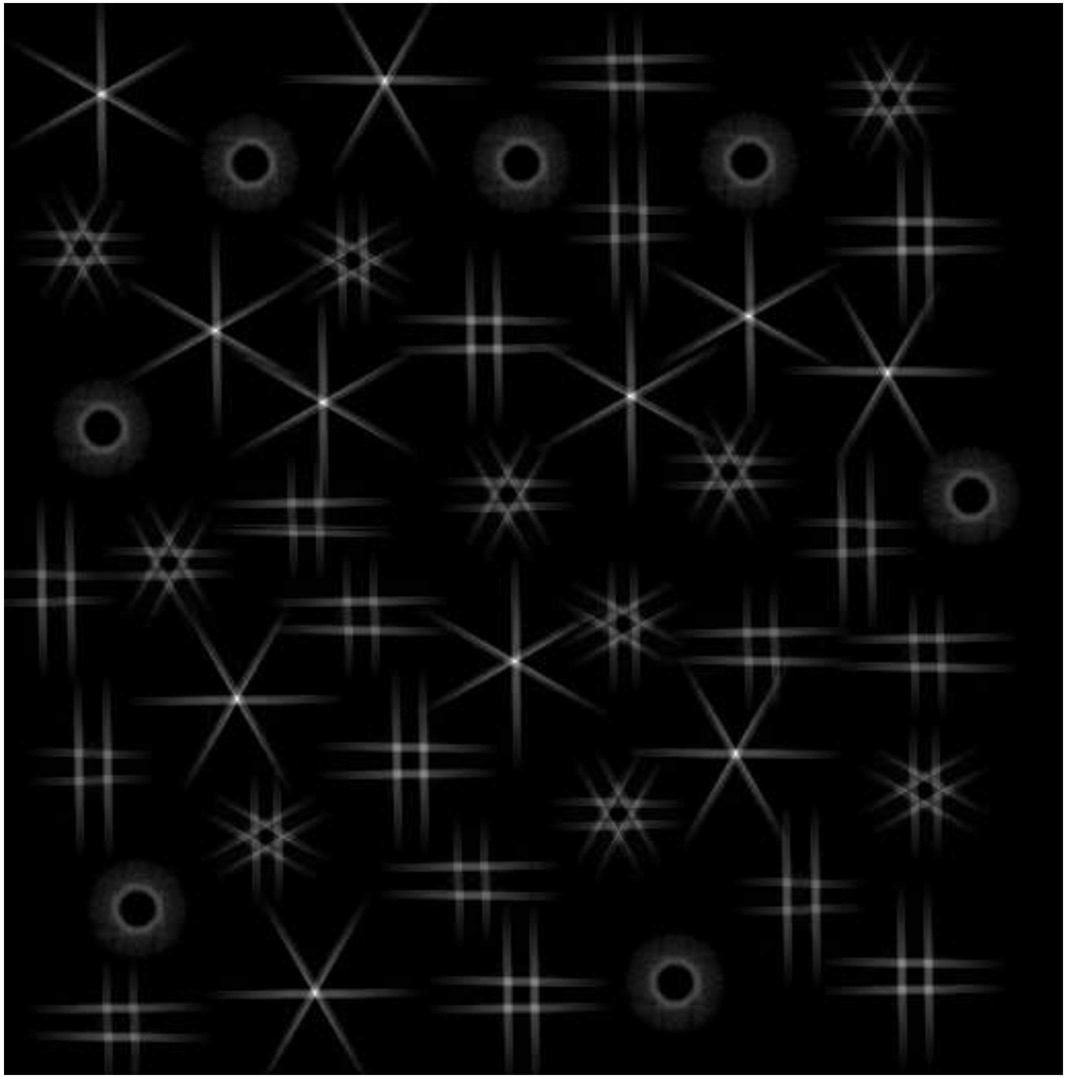


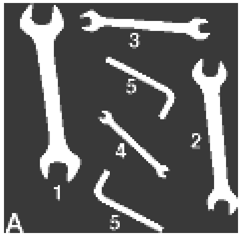




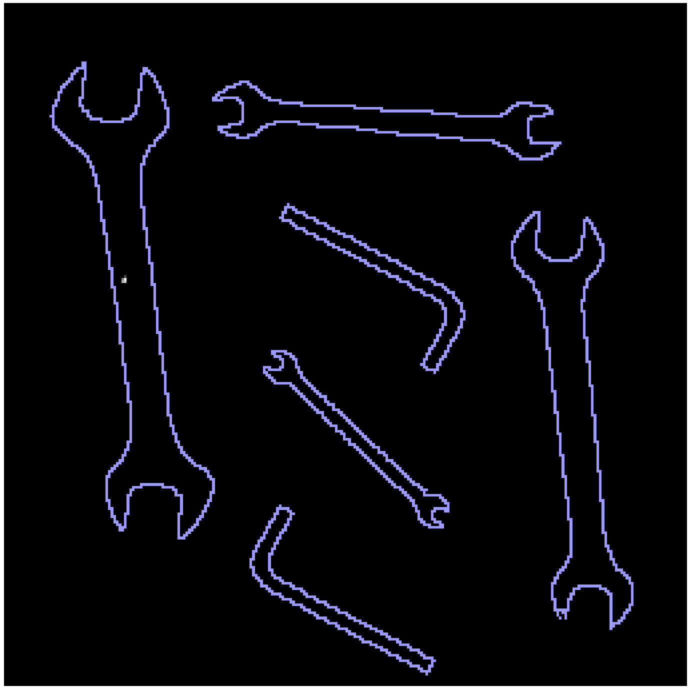
looking for a triangle

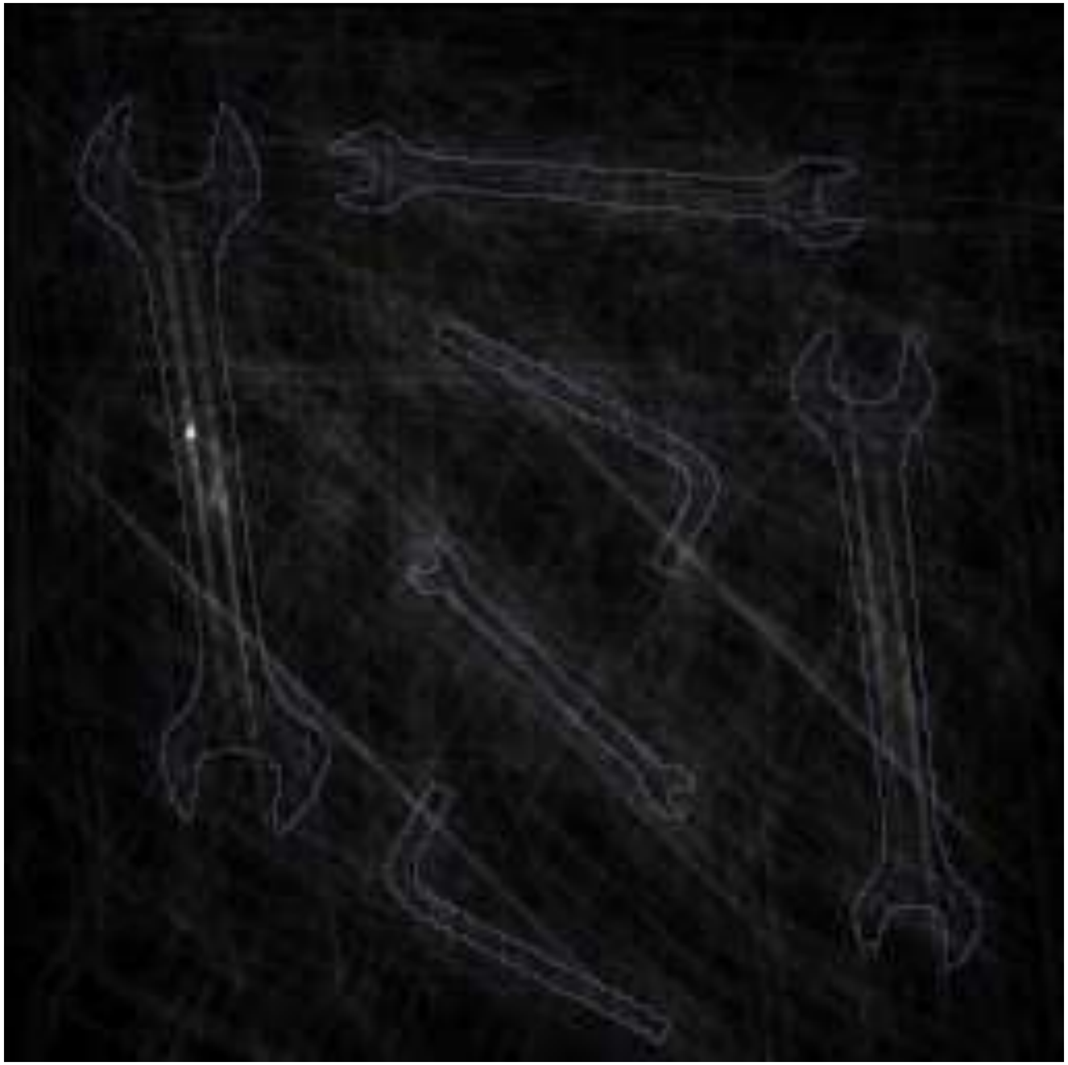


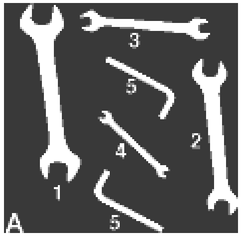




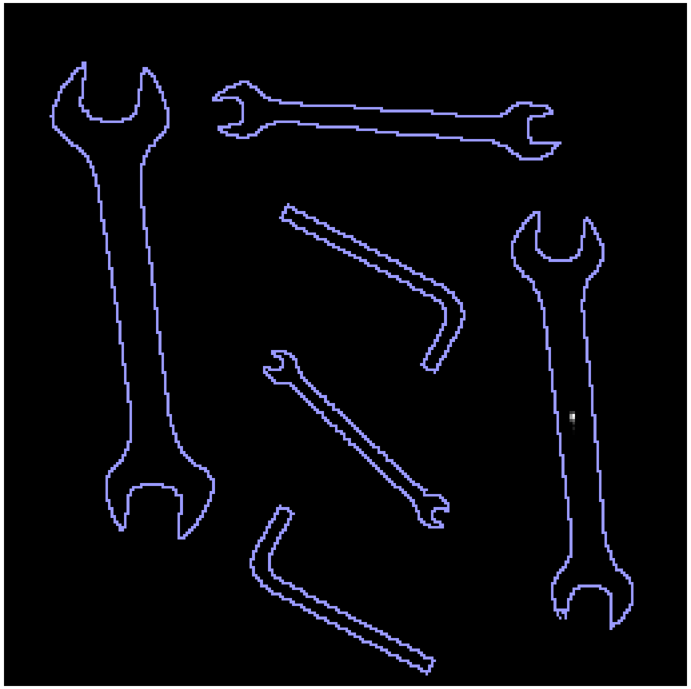
Example: wrench 1

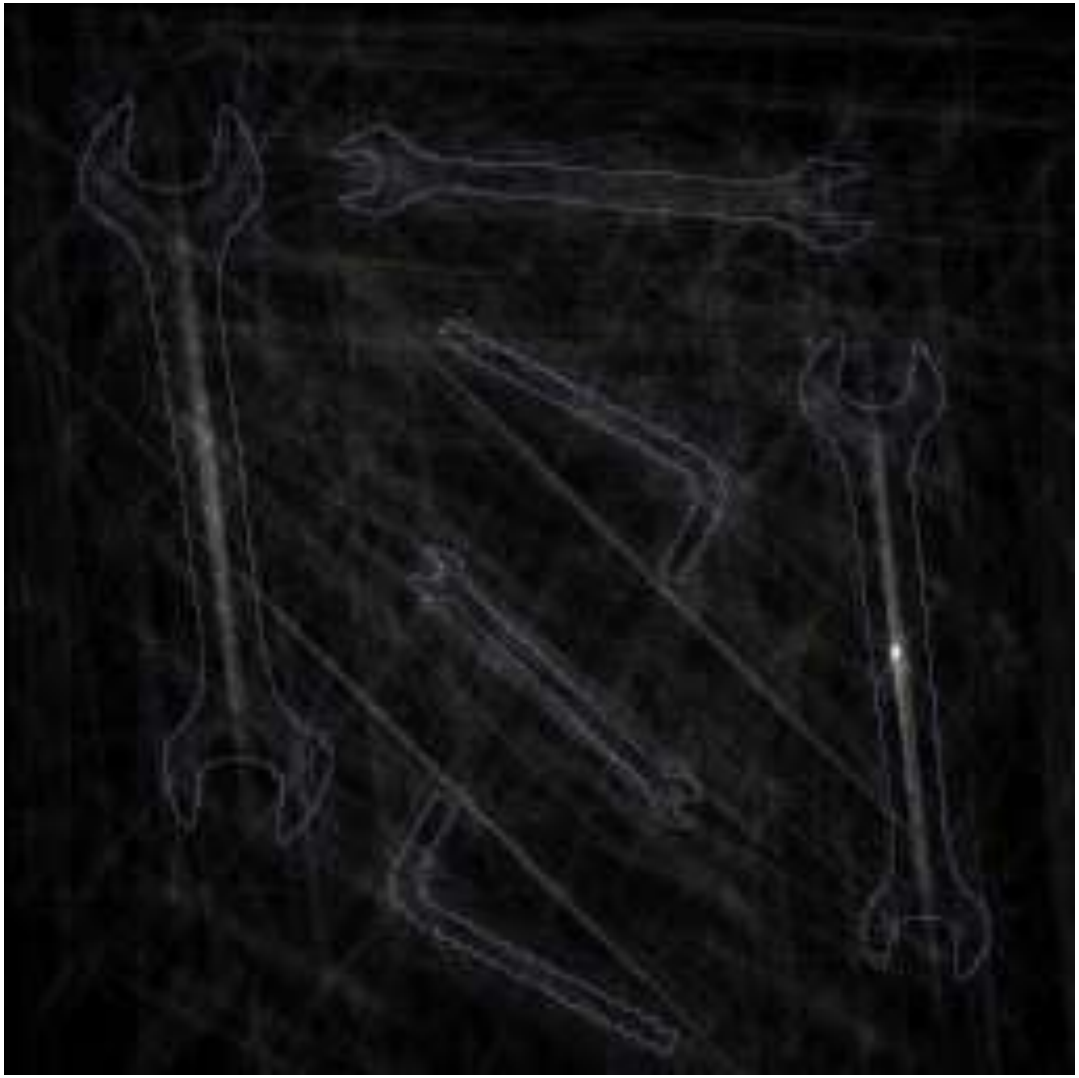


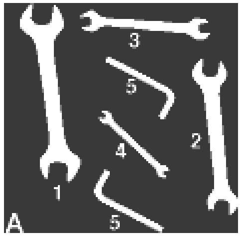




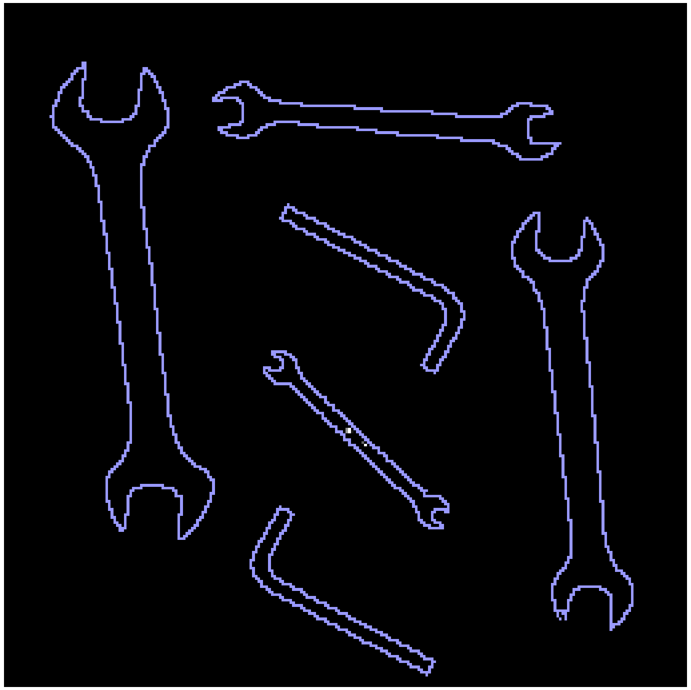
Example: wrench 2

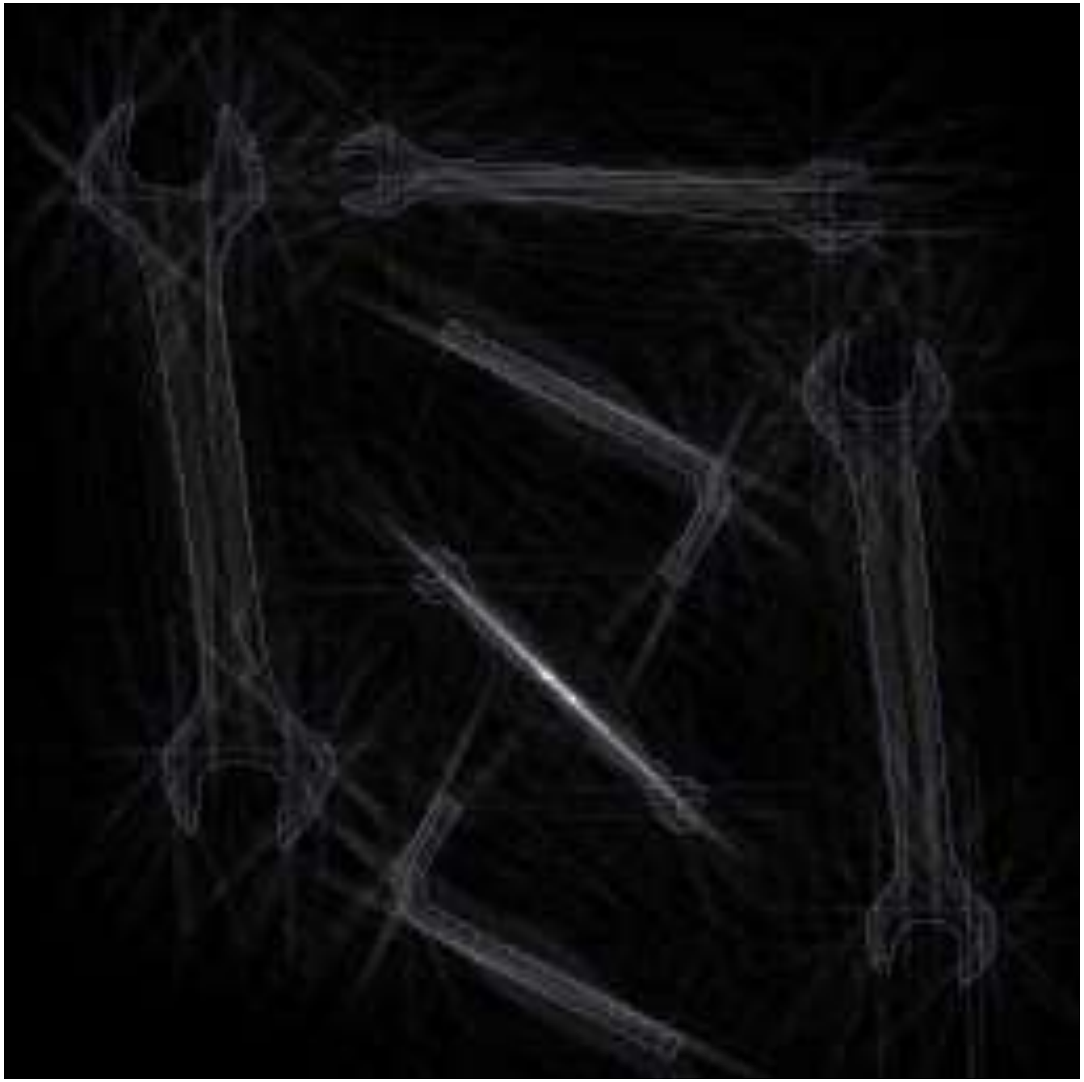


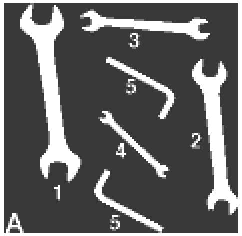




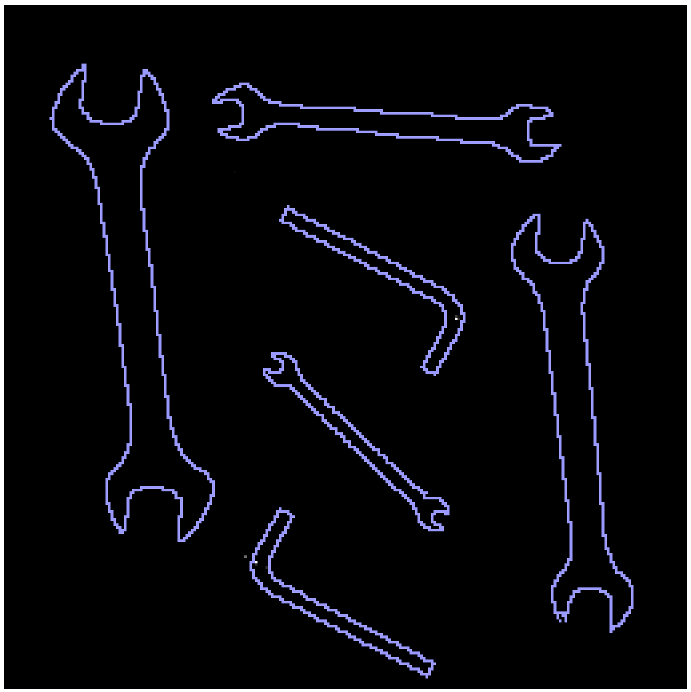
Example: wrench 4

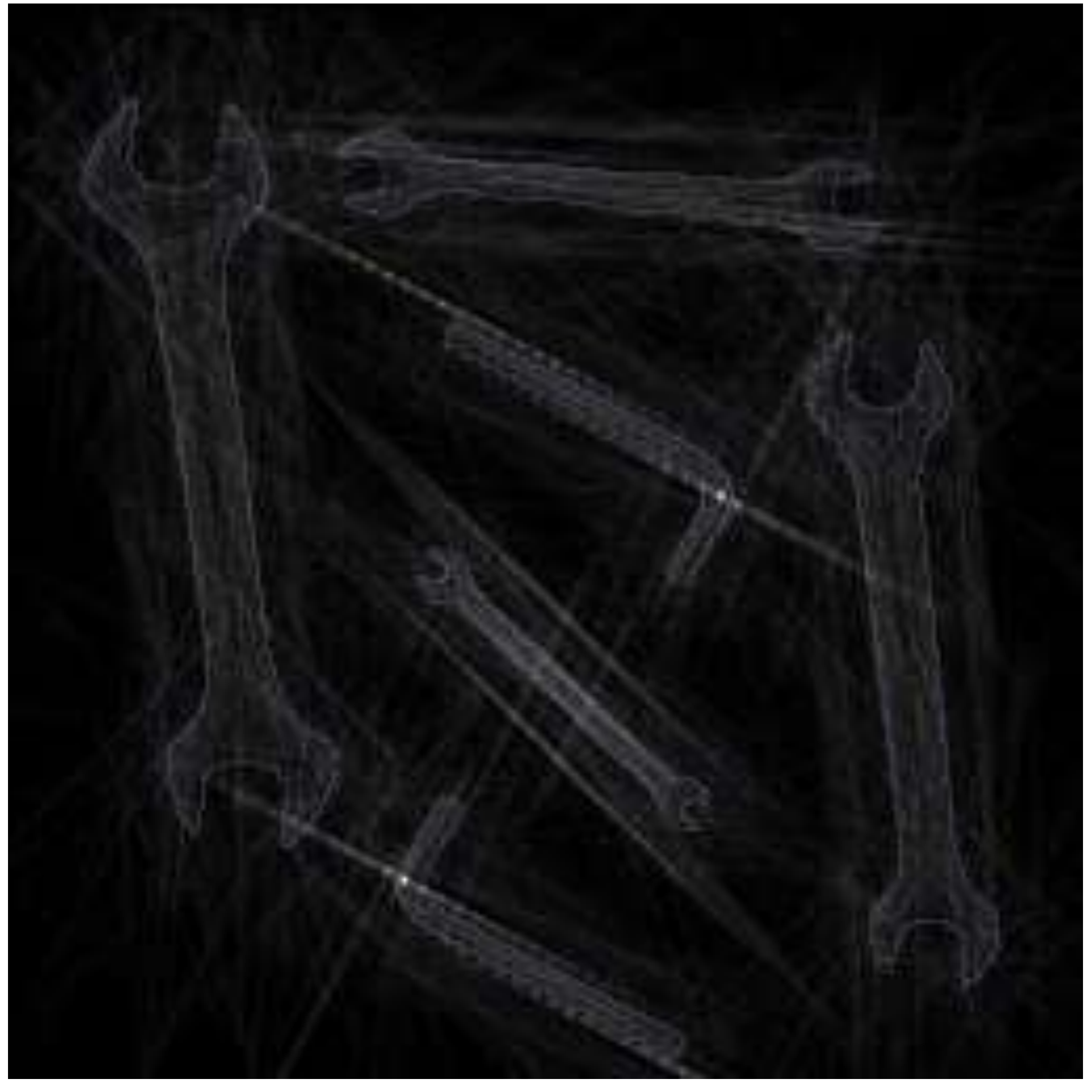




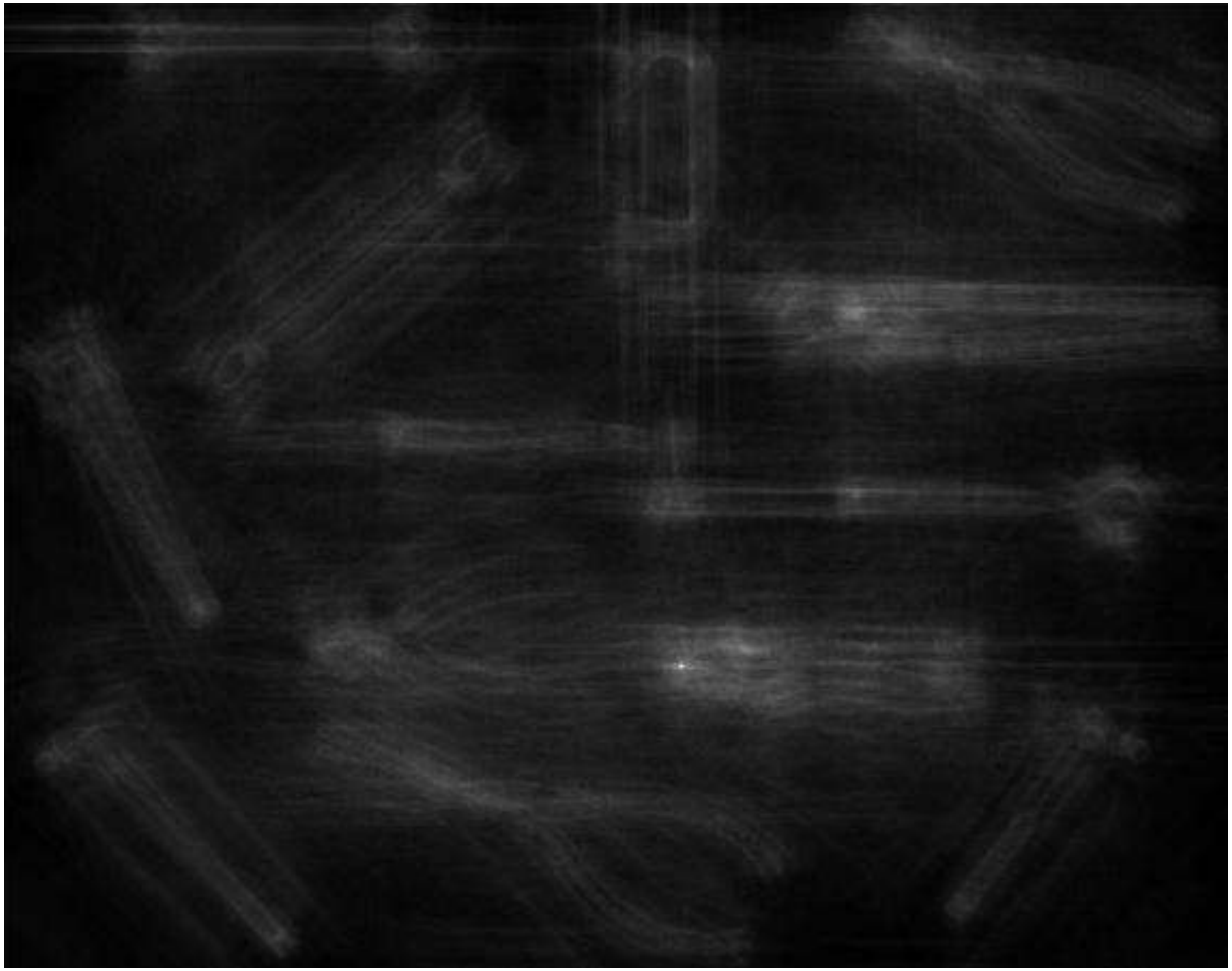


Example: hex key 5



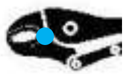


GHT: arbitrary pattern

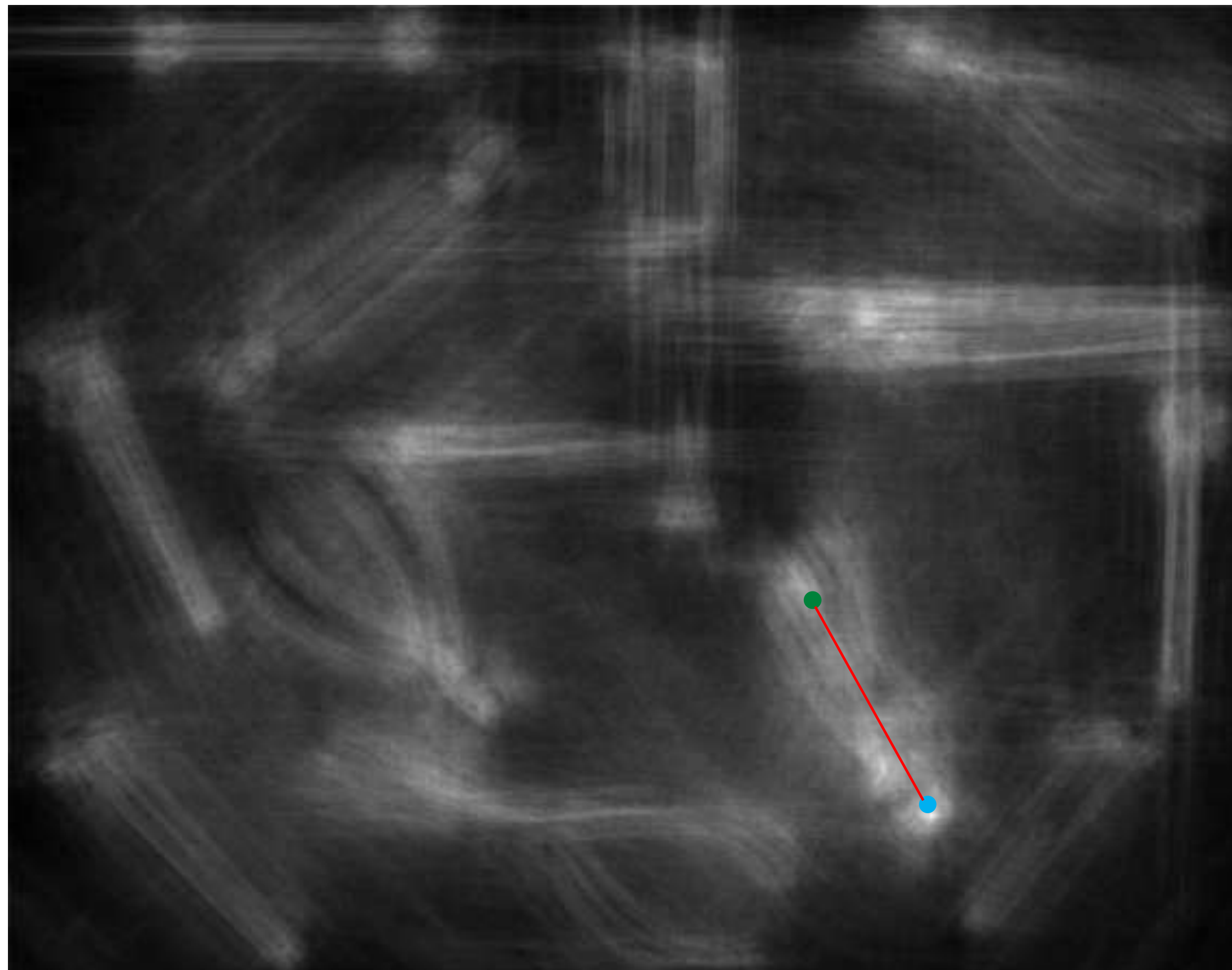


GHT: segmented pattern

ORIENTATION DETECTION

First pattern 

Second pattern 



Implementation aspects

- The RT can be decomposed in many subtable (maybe also overlapped) on the bases of **labels encoding some peculiarities** (e.g. a taxonomy of concavities and convexities)
- In the image plane, for each evidence, a corresponding sub-table is selected and **only this sub-table is involved in the voting process**
- **The intensity of the peak remains the same**, but it is reduced the number of scattered contributions: **so increasing the signal to noise ratio of the PS**

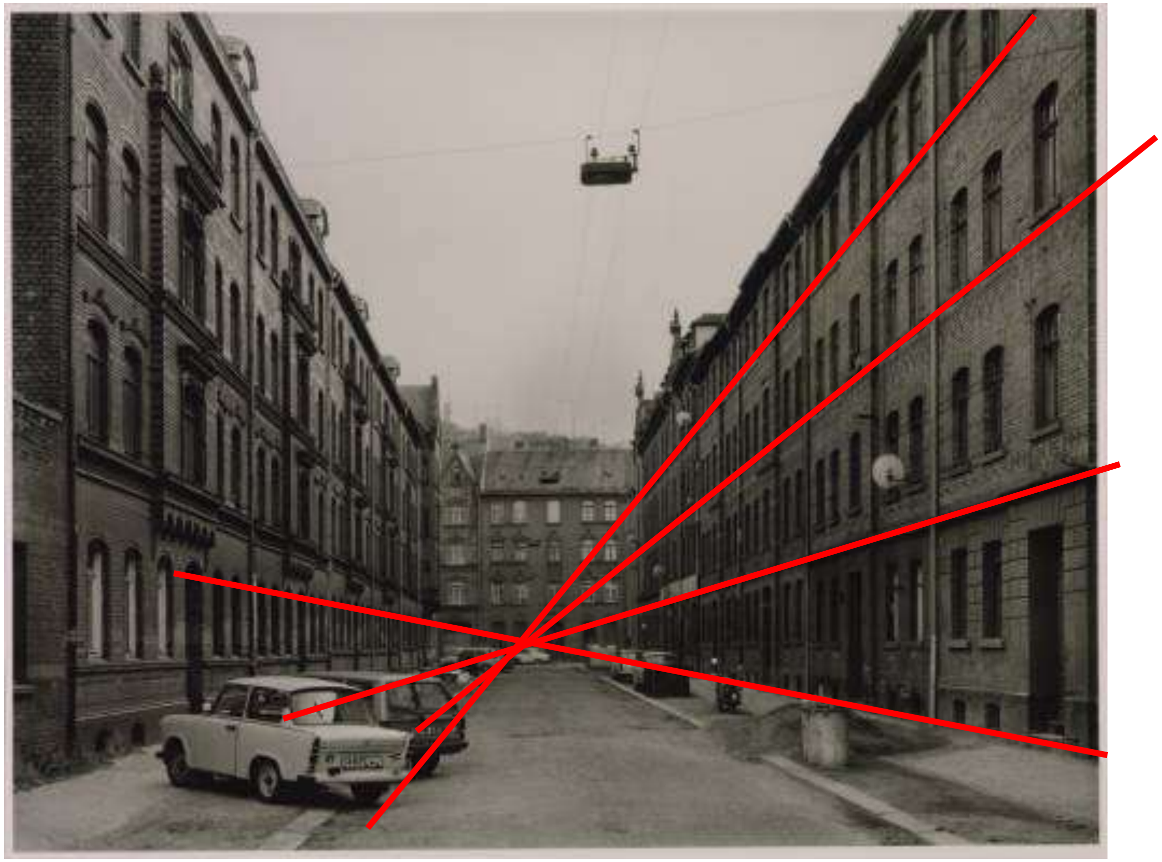
Computation performances

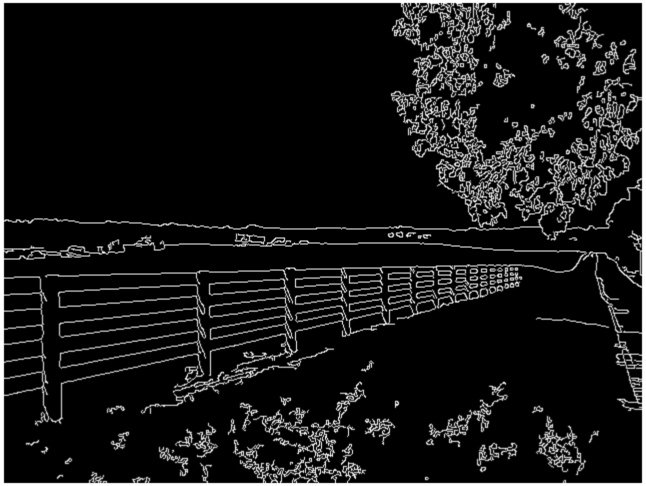
- The computation time is **linear with the product of the number of edge points in the image N_E with N_{RT}, the cardinality of the RT**
- In the case of labeled RT the computation time is given by **the weighted sum the sub-table cardinality by the number of occurrences of the correspondent labels**
- The algorithm is completely **parallelizable both over the image** (PEs taking care of different image blocks) and **over the RT** (PEs taking care of different object segment)

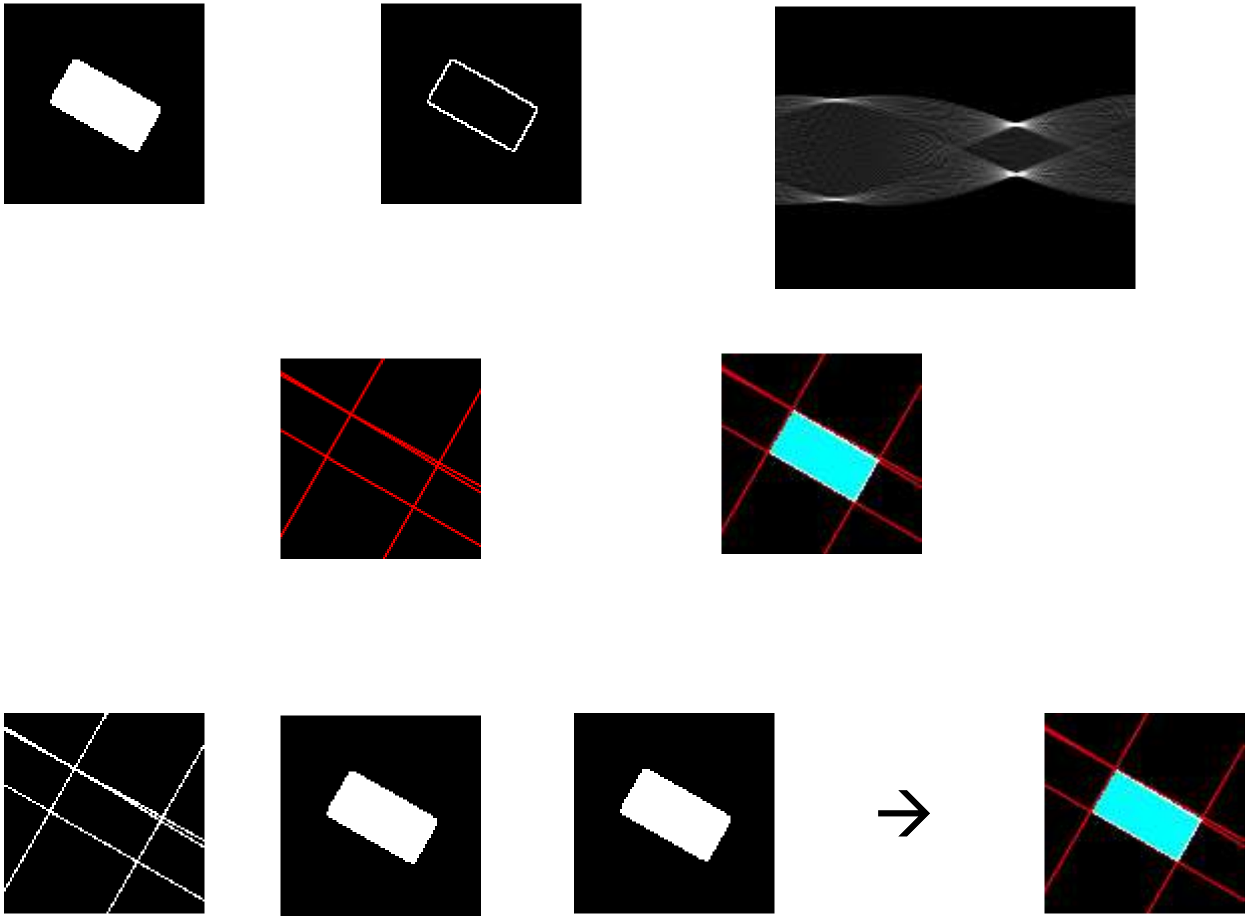
The Generalized Hough Transform

- For a given point P(x,y) obtained by an edge detector on the image under analysis, the set of compatible points of the PS (which here represent the positions of the of the P_{ref} compatible with the contour crossing P and the **PS coincide with the image space**) are computed.
- In case of rigid object and fixed size, for each contour point detected on the image, **a number of contributions onto PS are determined**. If all the contour points of a searched pattern, are effectively present in the image are detected properly, **a peak of value N will appear in the P_{ref} position of PS** (corresponding to the rigid motion model)
- Instead, considering the case of unknown scale factor **s**, to the image describing the position of the P_{ref} in the image, **an extra dimension must be introduced in the PS for the parameter s**. It become a 3D PS (replicating the image for each value of **s**) and in the above equation ρ must be multiplied by the correspondent scale factor **s**.
- Then, a local analysis, must be applied to evaluate the object orientation. Alternatively, the object orientation, for a 2D object, can be detect directly in a 4D PS: $(x_{ref}, y_{ref}, s, \varphi)$. **A convenient solution is to choose a couple of P_{ref}.**

Vanishing points and lines





Red channel Green channel Blue channel Color image