Region detectors
Requirements for region detection

• For region detection invariance transformations that should be considered are *illumination changes, translation, rotation, scale* and *full affine transform* (i.e. a region should correspond to the same pre-image for different viewpoints. Viewpoint changes can be locally approximated by affine transform if assuming locally planar objects and orthographic camera, that is perspective effects ignored)

• Region detection should be repeatable and stable, and capable to discriminate between regions
Affine Invariant Intensity Extrema–based

- The algorithm
  - Start from a local intensity extremum point
  - Go in every direction until the point of extremum of some function $f$. The curve connecting the points is the region boundary
  - Compute geometric moments of orders up to 2 for this region
  - Replace the region with ellipse
• Detecting extremal regions:
  – detect anchor points (f.e. using Harris detector for corners).
    Anchor points detected at multiple scales are local extremas of intensity
  – explore image around rays from each anchor point.
    Go along every ray starting from this point until an extremum of function $f$ is reached.

\[
f(t) = \frac{|I(t) - I_0|}{\frac{1}{t} \int_0^t |I(t) - I_0| \, dt}
\]

$f$ characteristic function of the region
(1 inside, 0 outside)

– all points create some irregularly-shaped region. Approximately corresponding regions are obtained for affine-transformed regions
Approximating the region

- Being function $f$ the characteristic function of the region, moments up to $2^{nd}$ order allow to approximate the region with an ellipse (the ellipse has the same $2^{nd}$ order moments as the region)

$$m_{pq} = \int x^p y^q f(x,y) \, dx \, dy$$

- The ellipse of an affine-transformed region corresponds to the ellipse of the original region under the same transformation

$$q = Ap$$

$$\Sigma_2 = A\Sigma_1 A^T$$

- As regions are illumination based warp ellipse to circle for affine invariance
MSER (Maximally Stable Extremal Regions)

- MSER is a method for blob detection in images

- MSER is based on the idea of taking regions which stay nearly the same through a wide range of thresholds. Regions are obtained by extracting the connecting components and then approximating with the bounding ellipse. Those regions descriptors are kept as features
**MSER (Maximally Stable Extremal Regions)**

- MSER is a method for blob detection in images. MSER is based on the idea of taking regions which stay nearly the same through a wide range of thresholds.

- The MSER algorithm extracts from an image a number of co-variant regions, called MSERs: an MSER is a *stable connected component of some gray-level sets of the image*.

- Optionally, elliptical frames are attached to the MSERs by fitting ellipses to the regions. Those regions descriptors are kept as features.

- The word “extremal” refers to the property that all pixels inside the MSER have either higher (bright extremal regions) or lower (dark extremal regions) intensity than all the pixels on its outer boundary.

The MSER extraction implements the following steps:
- Sweep threshold of intensity from black to white, performing a simple luminance thresholding of the image
- Extract connected components ("Extremal Regions")
- Find a threshold when an extremal region is "Maximally Stable", i.e. local minimum of the relative growth of its square
- Approximate a region with an ellipse (this step is optional)
- Keep those regions descriptors as features

\[ Q_{i^*} : i^* = \arg \min_i \left| Q_{i+\Delta} \setminus Q_{i-\Delta} \right| / |Q_i| \]
Sweeping image thresholds

From white to black

From black to white
• All the pixels below a threshold are white. The others are black

• Considering a sequence of thresholded images with increasing thresholds, sweeping from black to white we pass from a black image to images where white blobs appear and grow larger by merging, up to the final image.

• The set of all connected components in the sequence is the set of maximal regions

• Over a large range of thresholds the local binarization is stable and shows some invariance to affine transformation of image intensities and scaling
• Due to the discrete nature of the image, the region below / above may be coincident with the actual region, in which case the region is still deemed maximal.

• However, even if an extremal region is maximally stable, it might be rejected if:
  – it is too big (there is a parameter MaxArea);
  – it is too small (there is a parameter MinArea);
  – it is too unstable (there is a parameter MaxVariation);
  – it is too similar to its parent MSER
MSER examples

Example 1
Threshold 1
Threshold 2
Threshold 3
Ears and Square bounded by ellipses
Example 2

Extracted MSERs for both bright-on-dark (green) and dark-on-bright (yellow).
MSER features

- MSER performs well on images containing homogeneous regions with distinctive boundaries.
- MSER works well for small regions
- MSER doesn’t work well with images with any motion blur
- Good repeatability
- Affine invariant
- A smart implementation makes it one of the fastest region detectors
MSER against other region detectors

• **View point change**: MSER performs the best.
• **Scale change**: Hessian-Affine, MSER, Harris-Affine are the best
• **Blur**: all detectors are invariant to image blur. MSER not invariant with any motion blur
• **JPEG change**: MSER is the best