Professor Brendan Morris, SEB 3216, brendan.morris@unlv.edu

ECG782: Multidimensional Digital Signal Processing

http://www.ee.unlv.edu/~b1morris/ecg782/

Outline

- Interest Point Detection
- Maximally Stable Regions

Detection of Corners (Interest Points)

- Useful for fundamental vision techniques
  - Image matching or registration
- Correspondence problem needs to find all pairs of matching pixels
  - Typically a complex problem
  - Can be made easier only considering a subset of points
- Interest points are these important image regions that satisfy some local property
  - Corners are a way to get to interest points

Feature Detection and Matching

- Essential component of modern computer vision
  - E.g. alignment for image stitching, correspondences for 3D model construction, object detection, stereo, etc.

- Need to establish some features that can be detected and matched

Determining Features to Match

- What can help establish correspondences between images?




Different Types of Features

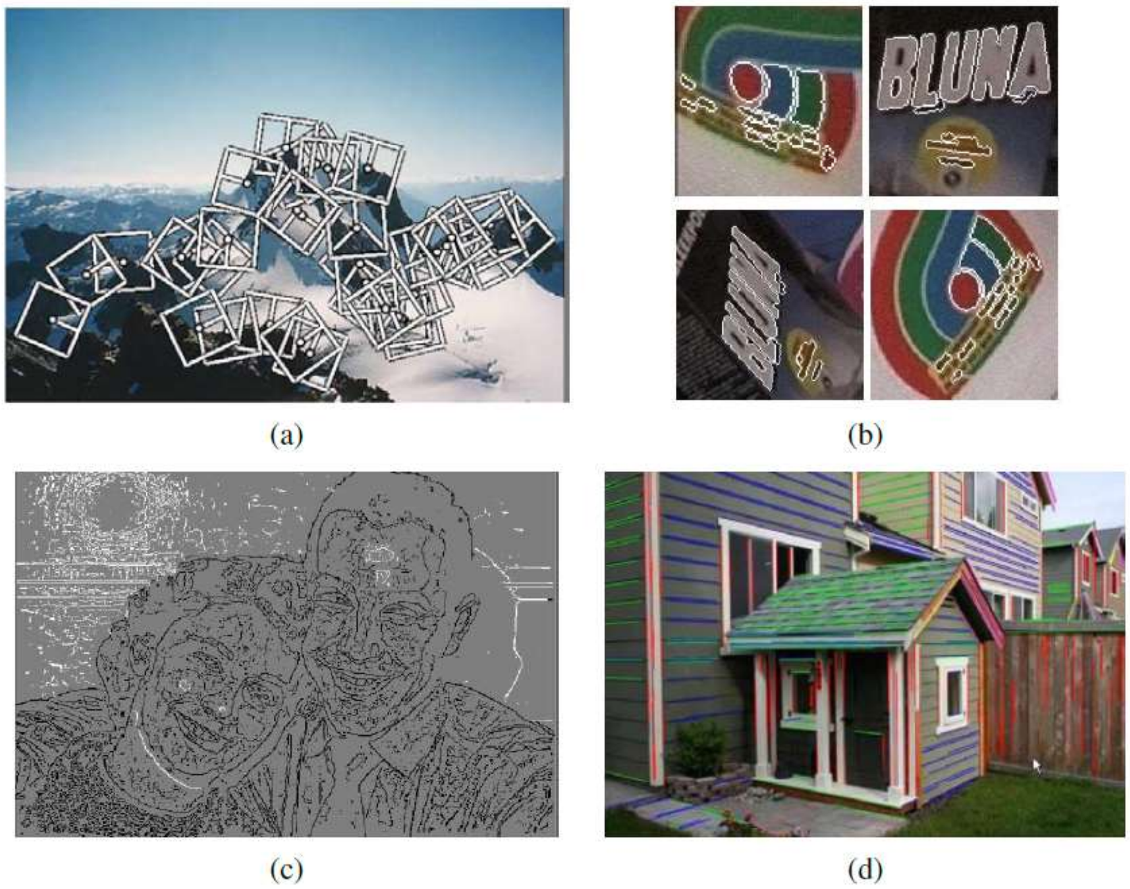


Figure 4.1 A variety of feature detectors and descriptors can be used to analyze, describe and match images: (a) point-like interest operators (Brown, Szeliski, and Winder 2005) © 2005 IEEE; (b) region-like interest operators (Matas, Chum, Urban *et al.* 2004) © 2004 Elsevier; (c) edges (Elder and Goldberg 2001) © 2001 IEEE; (d) straight lines (Sinha, Steedly, Szeliski *et al.* 2008) © 2008 ACM.

Different Types of Features

- Points and patches
- Edges
- Lines

- Which features are best?
  - Depends on the application
  - Want features that are robust
    - Descriptive and consistent (can readily detect)

Points and Patches

- Maybe most generally useful feature for matching
  - E.g. Camera pose estimation, dense stereo, image stitching, video stabilization, tracking
  - Object detection/recognition

- Key advantages:
  - Matching is possible even in the presence of clutter (occlusion)
  - and large scale and orientation changes

Point Correspondence Techniques

- Detection and tracking
  - Initialize by detecting features in a single image
  - Track features through localized search
  - Best for images from similar viewpoint or video
- Detection and matching
  - Detect features in all images
  - Match features across images based on local appearance
  - Best for large motion or appearance change

Keypoint Pipeline

- Feature detection (extraction)
  - Search for image locations that are likely to be matched in other images
- Feature description
  - Regions around a keypoint are represented as a compact and stable descriptor
- Feature matching
  - Descriptors are compared between images efficiently
- Feature tracking
  - Search for descriptors in small neighborhood
  - Alternative to matching stage best suited for video

Feature Detectors

- Must determine image locations that can be reliably located in another image

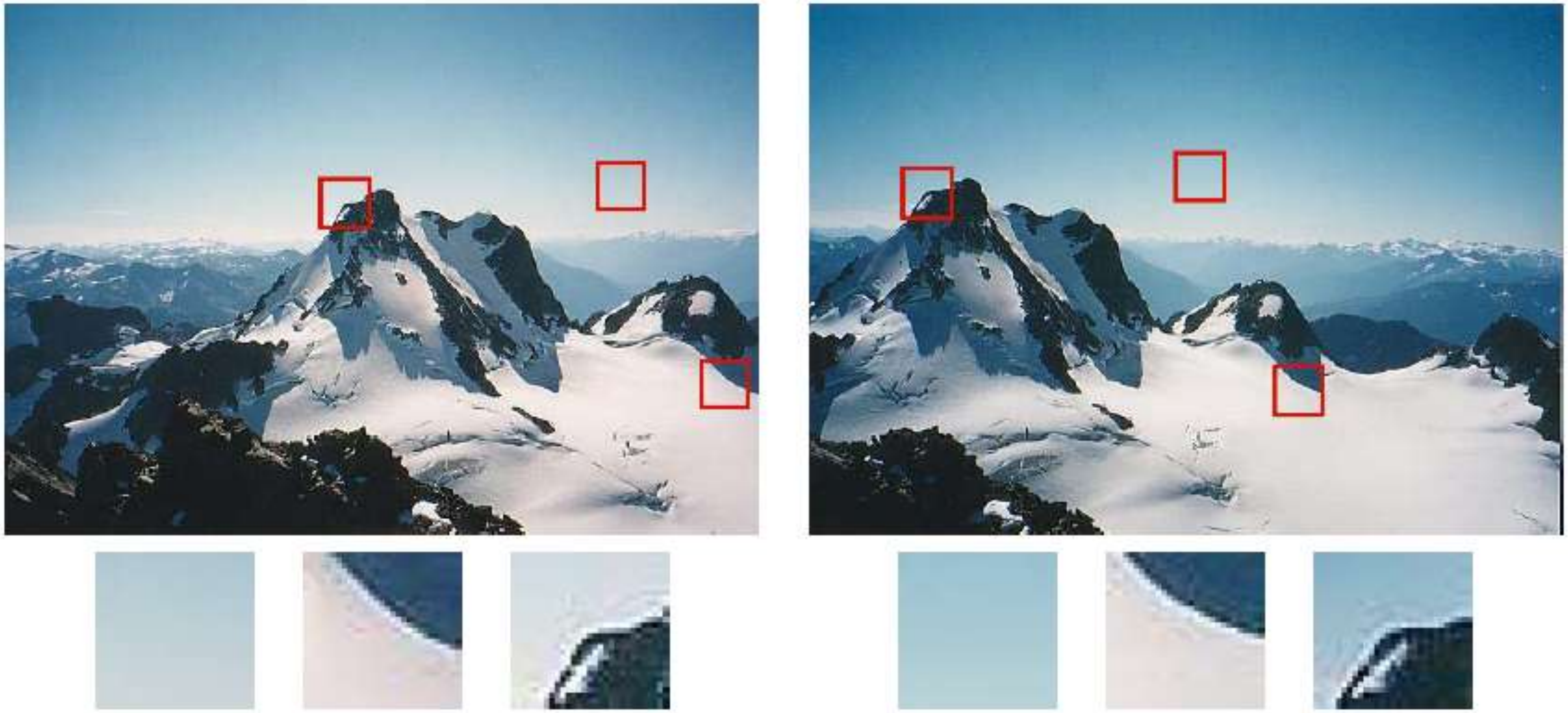


Figure 4.3 Image pairs with extracted patches below. Notice how some patches can be localized or matched with higher accuracy than others.

Comparison of Image Patches

- Textureless patches
  - Nearly impossible to localize and match
    - Sky region "matches" to all other sky areas

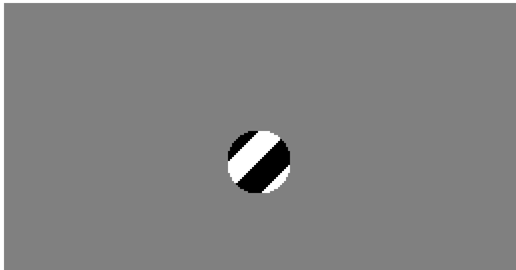
- Edge patches
  - Large contrast change (gradient)
  - Suffer from aperture problem
    - Only possible to align patches along the direction normal the edge direction

- Corner patches
  - Contrast change in at least two different orientations
  - Easiest to localize

Aperture Problem I

- Only consider a small window of an image
  - Local view does not give global structure
  - Causes ambiguity

- Best visualized with motion (optical flow later)
  - Imagine seeing the world through a straw hole
  - Aperture Problem - MIT – Demo
  - Also known as the barber pole effect



Source: Wikipedia

Aperture Problem II

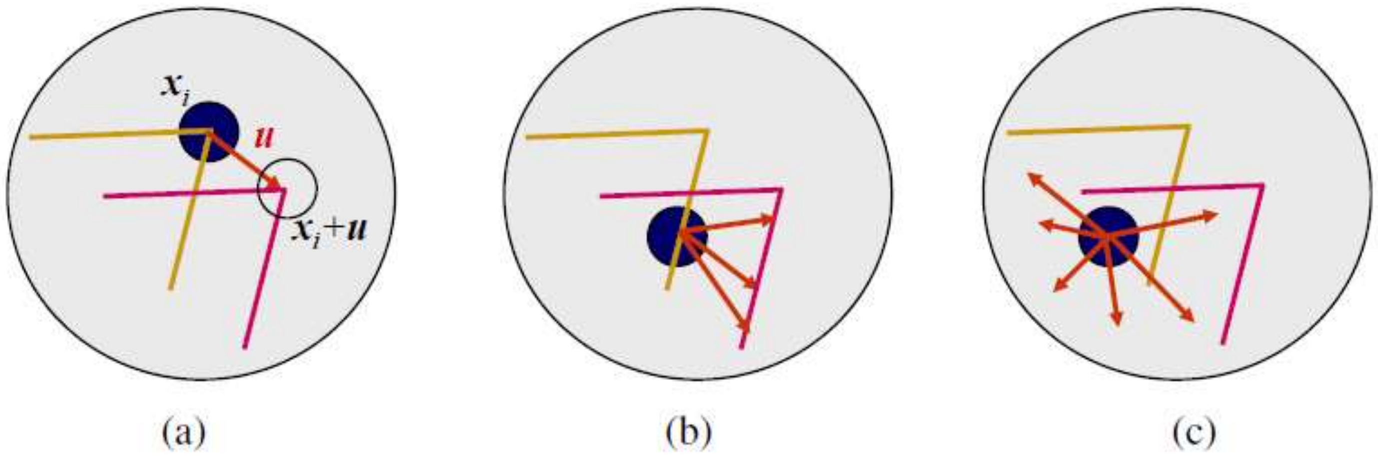


Figure 4.4 Aperture problems for different image patches: (a) stable ("corner-like") flow; (b) classic aperture problem (barber-pole illusion); (c) textureless region. The two images I_0 (yellow) and I_1 (red) are overlaid. The red vector $\boldsymbol{u}$ indicates the displacement between the patch centers and the $w(\boldsymbol{x}_i)$ weighting function (patch window) is shown as a dark circle.

- Corners have strong matches
- Edges can have many potential matches
  - Constrained upon a line
- Textureless regions provide no useful information

WSSD Matching Criterion

- Weighted summed squared difference
  - $E_{WSSD}(\boldsymbol{u}) = \sum_i w(\boldsymbol{x}_i)\,[I_1(\boldsymbol{x}_i - \boldsymbol{u}) - I_0(\boldsymbol{x}_i)]^2$
    - I_1, I_0 - two image patches to compare
    - $\boldsymbol{u} = (u, v)$ – displacement vector
    - $w(\boldsymbol{x})$ - spatial weighting function
- Normally we do not know the image locations to perform the match
  - Calculate the autocorrelation in small displacements of a single image
    - Gives a measure of stability of patch
  - $E_{AC}(\Delta\boldsymbol{u}) = \sum_i w(\boldsymbol{x}_i)\,[I_0(\boldsymbol{x}_i - \Delta\boldsymbol{u}) - I_0(\boldsymbol{x}_i)]^2$

Image Patch Autocorrelation

$$E_{AC}(\Delta \boldsymbol{u}) = \sum_i w(\boldsymbol{x}_i)\,[I_0(\boldsymbol{x}_i - \Delta \boldsymbol{u}) - I_0(\boldsymbol{x}_i)]^2$$
$$= \sum_i w(\boldsymbol{x}_i)\,[\nabla I_0(\boldsymbol{x}_i) \cdot \Delta \boldsymbol{u}]^2$$
$$= \Delta \boldsymbol{u}^T A \Delta \boldsymbol{u}$$

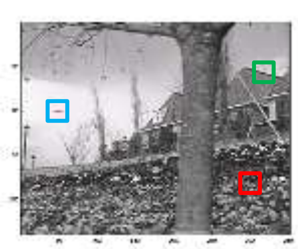
- $\nabla I_0(\boldsymbol{x}_i)$ - image gradient
  - We have seen how to compute this
- A – autocorrelation matrix

$$A = w * \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix}$$

  - Compute gradient images and convolve with weight function
  - Also known as second moment matrix
  - (Harris matrix)

- Example autocorrelation

Image Autocorrelation II



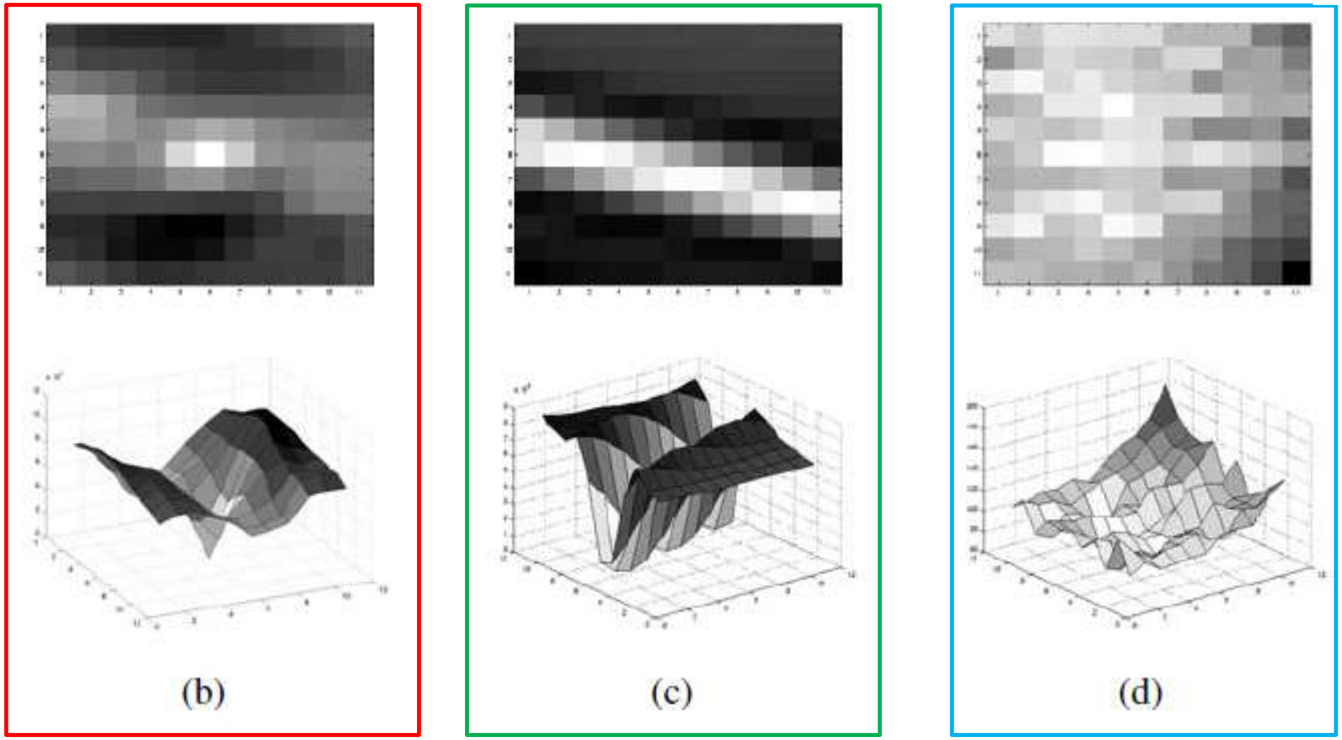
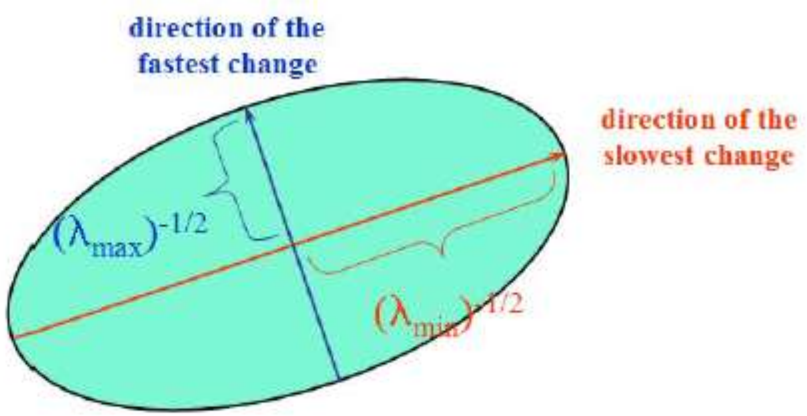


Figure 4.5 Three auto-correlation surfaces $E_{\mathrm{AC}}(\Delta u)$ shown as both grayscale images and surface plots: (a) The original image is marked with three red crosses to denote where the auto-correlation surfaces were computed; (b) this patch is from the flower bed (good unique minimum); (c) this patch is from the roof edge (one-dimensional aperture problem); and (d) this patch is from the cloud (no good peak). Each grid point in figures b–d is one value of Δu.

Image Autocorrelation III

- The matrix *A* provides a measure of uncertainty in location of the patch
- Do eigenvalue decomposition
  - Get eigenvalues and eigenvector directions
- Good features have both eigenvalues large
  - Indicates gradients in orthogonal directions (e.g. a corner)

- Uncertainty ellipse



- Many different methods to quantify uncertainty
  - Easiest: look for maxima in the smaller eigenvalue

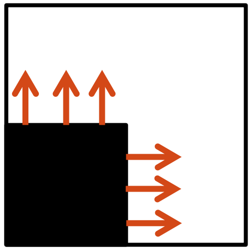
Basic Feature Detection Algorithm

1. Compute the horizontal and vertical derivatives of the image I_x and I_y by convolving the original image with derivatives of Gaussians (Section 3.2.3).
2. Compute the three images corresponding to the outer products of these gradients. (The matrix $\boldsymbol{A}$ is symmetric, so only three entries are needed.)
3. Convolve each of these images with a larger Gaussian.
4. Compute a scalar interest measure using one of the formulas discussed above.
5. Find local maxima above a certain threshold and report them as detected feature point locations.

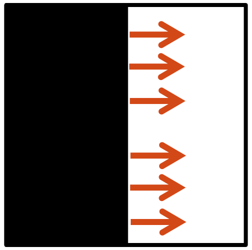
Algorithm 4.1 Outline of a basic feature detection algorithm.

Interest Point Detection

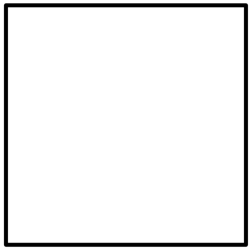
- The correlation matrix gives a measure of edges in a patch
- Corner
  - Gradient directions
    - $\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$
  - Correlation matrix
    - $A \propto \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$



- Edge
  - Gradient directions
    - $\begin{bmatrix}1\\0\end{bmatrix}$
  - Correlation matrix
    - $A \propto \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$



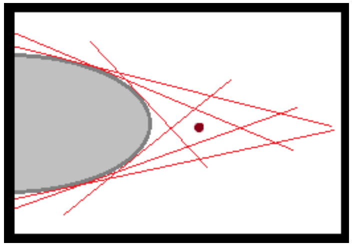
- Constant
  - Gradient directions
    - $\begin{bmatrix}0\\0\end{bmatrix}$
  - Correlation matrix
    - $A \propto \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$



Harris Corners

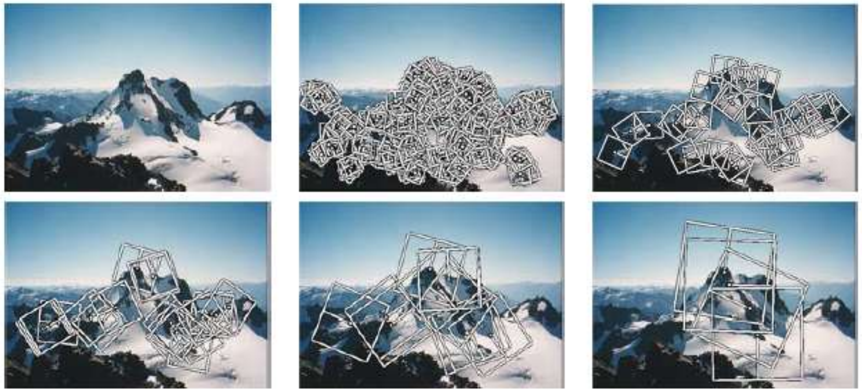
Improving Feature Detection

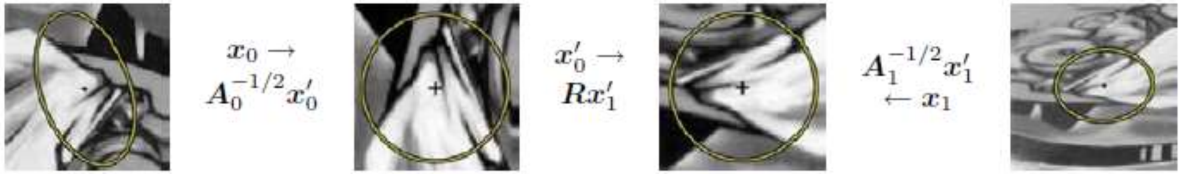
- Corners may produce more than one strong response (due to neighborhood)
  - Estimate corner with subpixel accuracy – use edge tangents
  - Non-maximal suppression – only select features that are far enough away
    - Create more uniform distribution – can be done through blocking as well
- Scale invariance
  - Use an image pyramid – useful for images of same scale
  - Compute Hessian of difference of Gaussian (DoG) image
  - Analyze scale space [SIFT – Lowe 2004]
- Rotational invariance
  - Need to estimate the orientation of the feature by examining gradient information
- Affine invariance
  - Closer to appearance change due to perspective distortion
  - Fit ellipse to autocorrelation matrix and use it as an affine coordinate frame
  - Maximally stable region (MSER) [Matas 2004] – regions that do not change much through thresholding



(a) Strongest 250

(c) ANMS 250, $r = 24$

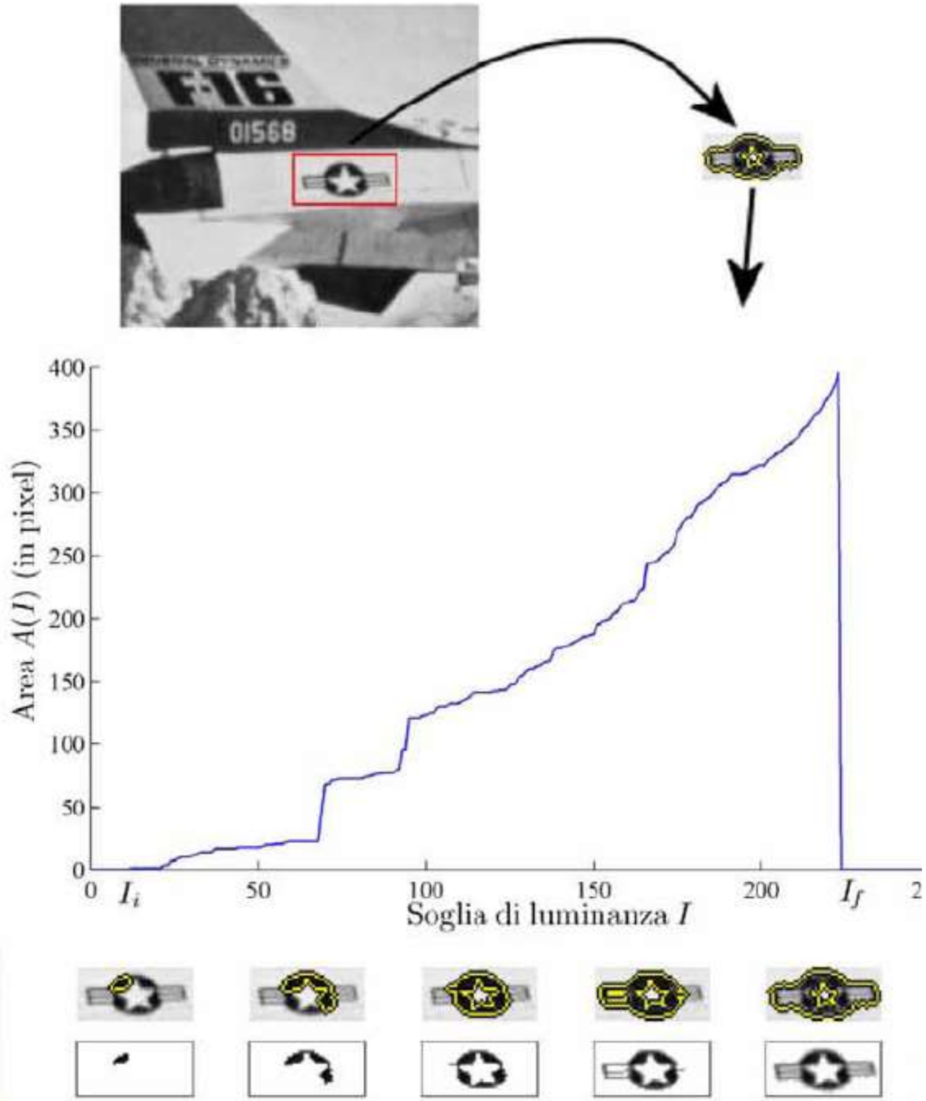




Maximally Stable Extremal Regions

- MSERs are image structures that can be recovered after translations, rotations, similarity (scale), and affine (shear) transforms
- Connected areas characterized by almost uniform intensity, surrounded by contrasting background
- Constructed based on a watershed-type segmentation
  - Threshold image a multiple different values
  - MSERs are regions with shape that does not change much over thresholds
- Each region is a connected component but no global or optimal threshold is selected

MSER



- Red borders from increasing intensity
- Green boarders from decreasing intensity

MSER Invariance

- Fit ellipse to area and normalize into circle