Image segmentation

Václav Hlaváč

Czech Technical University in Prague

Czech Institute of Informatics, Robotics and Cybernetics

160 00 Prague 6, Jugoslávských partyzánů 1580/3, Czech Republic

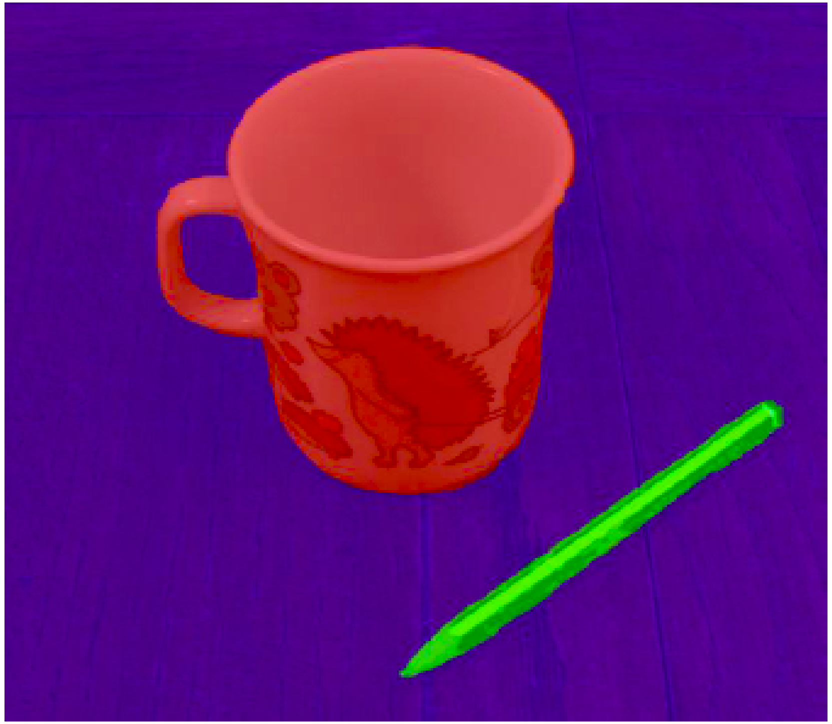
http://people.ciirc.cvut.cz/hlavac, vaclav.hlavac@cvut.cz

also Center for Machine Perception, http://cmp.felk.cvut.cz

Outline of the talk:

- What is segmentation? Segmentation is application dependent because it needs image interpretation.
- Taxonomy of segmentation methods.
- Thresholding-based segmentation.
- . . . *the rest comes in another presentation.*

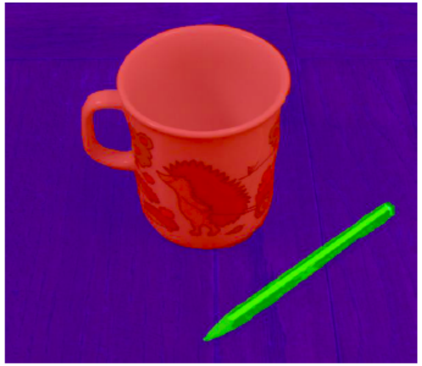
What is segmentation? Motivating picture



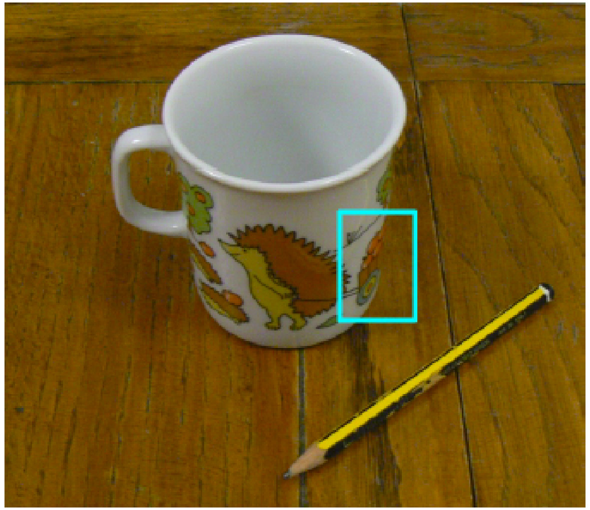
Image, courtesy Ondřej Drbohlav

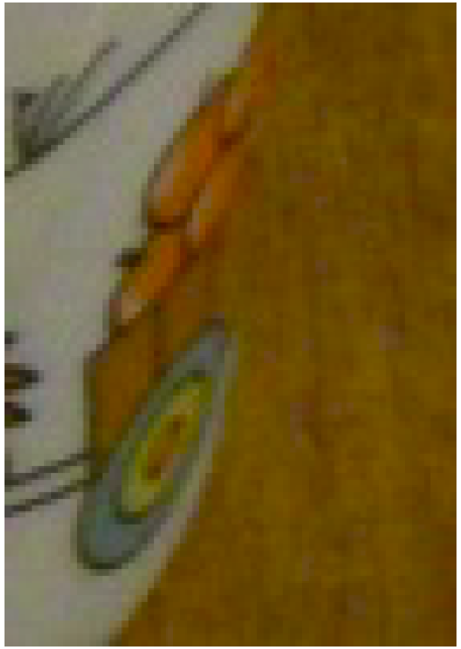
What is image segmentation ?

- Segmentation is a collection of methods allowing to interpret spatially close parts of the image as objects.
- Regions (i.e., compact sets) represent spatial closeness naturally and thus are important building steps towards segmentation. Objects in a 2D image very often correspond to distinguishable regions.
- The object is everything what is of interest in the image (from the particular application point of view). The rest of the image is background.
- The approach is similar to that used in pattern recognition, i.e., division of the image into set of equivalence classes.



Segmentation can be difficult, example



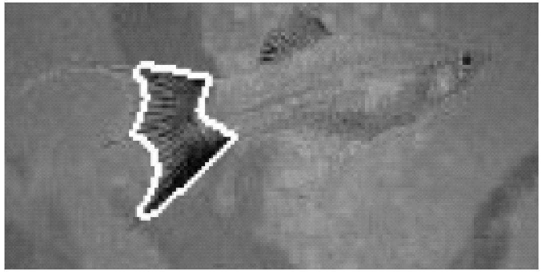


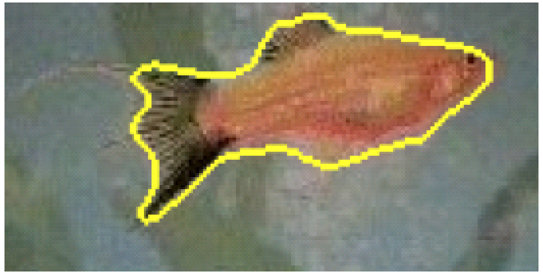
Image, courtesy Ondřej Drbohlav

- It is difficult to find border between the cup and background in the indicated region because it does not differ in a local view.
- Only knowledge of the cup semantics can solve the puzzle.

Image segmentation, a bit of magic

- There is often no single answer how to segment.
- Segmentation is mostly based on rather ad hoc methods.
- There is no encompassing broad theory of segmentation. However, several recent theoretically grounded approaches have formulated segmentation as an optimization task (e.g., in a Markovian fields formalism).
- The special case of foreground vs. background segmentation is often met.
- Segmentation usually makes sense in a scope of a particular application.





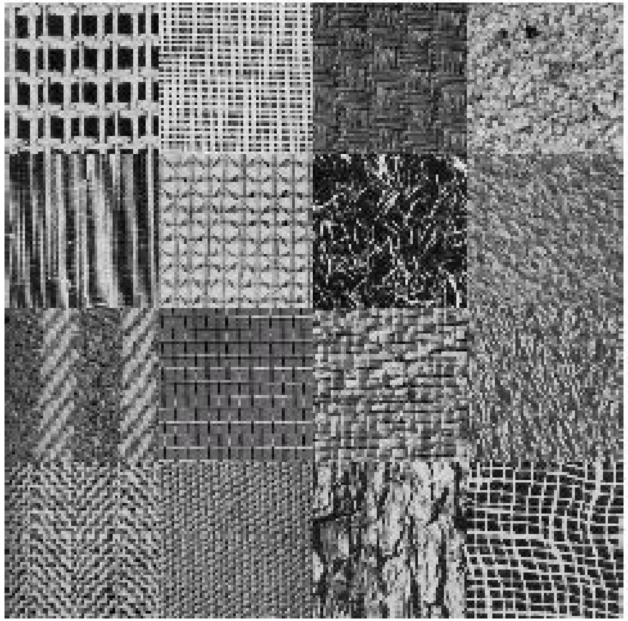
Courtesy, images: Thomas Brox, TU Dresden, 2008

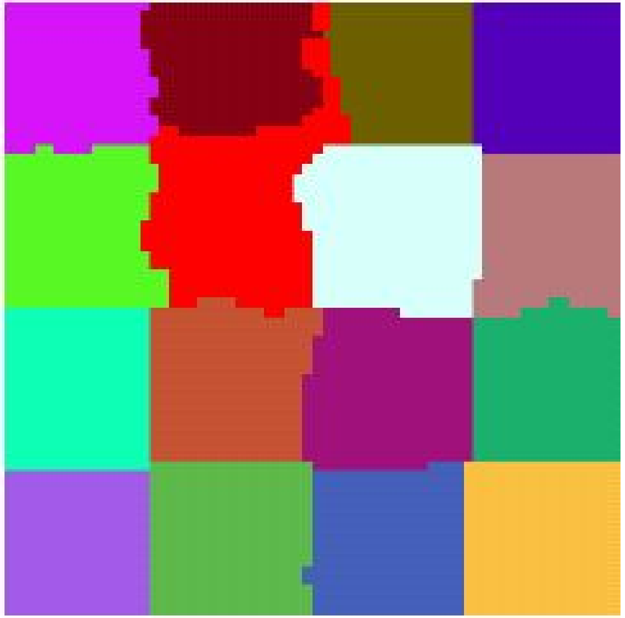
Segmentation is application dependent

- Segmentation depends on an application, its semantics.
- Methods are not universally applicable to all images.
- Direct segmentation of the input image takes pragmatically into account semantic information about a particular application.
- This is the way how to bypass ill-posed problems in inverted tasks related to image formation physics.
- A vital role of a priori information:
  - Low-level: e.g., brightness, spatial coherence, color, texture, motion, ...
  - Mid-level: object symmetries, proximity on larger scale, ...

Segmentation, Example 1

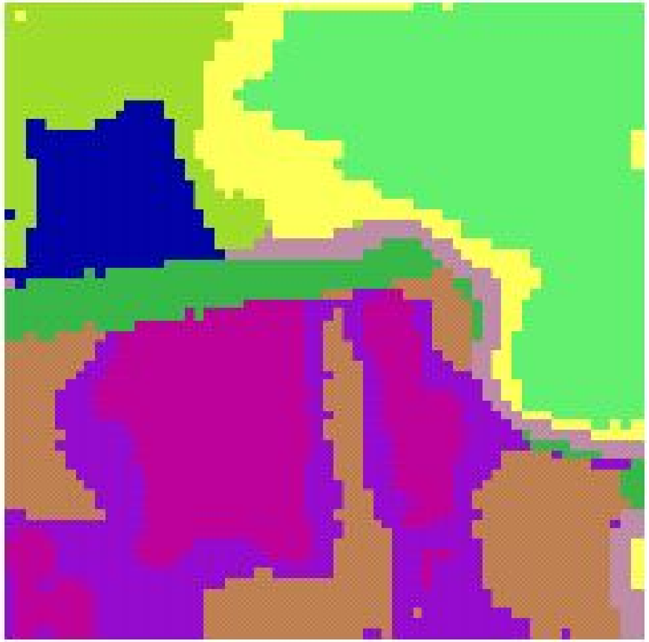
Segmentation according to the texture.





Segmentation, Example 2

Segmentation of an aerial image of the sea coast.



Complete vs. partial segmentation (1)

Complete segmentation -- divides an image into non-overlapping regions that match to the real world objects.

Complete segmentation divides an image R into the finite number S of regions $R_1, \ldots, R_S$

$$R = \bigcup_{i=1}^{S} R_i \,, \qquad R_i \cap R_j = \emptyset \,, \qquad i \neq j \,.$$

Partial segmentation -- it is possible to find only parts with semantic meaning in the image (e.g., regions, collection of edgels) which will lead to interpretation in later analysis.

Complete vs. partial segmentation (2)

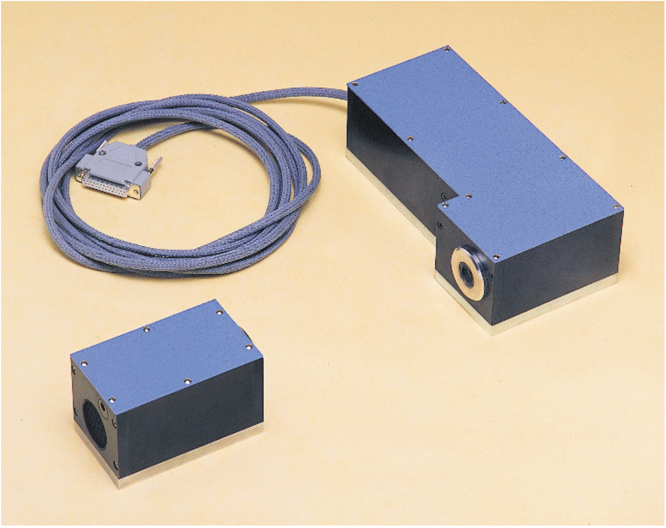
- To proceed from partial to complete segmentation it is needed to explore the higher-level of information processing.
- This can be performed iteratively in a feed-back loop.
- The information about the semantics of the specific application is used.
- Partial segmentation reduces the amount of data which need to be processed.

Examples of a complete 2D segmentation

Seek contrast objects in a homogeneous background. Intensity thresholding provides a silhouette corresponding to objects, *e.g., printed characters, cell kernels, back-light illuminated details inspected in industry.*

Example – A digital profile projector

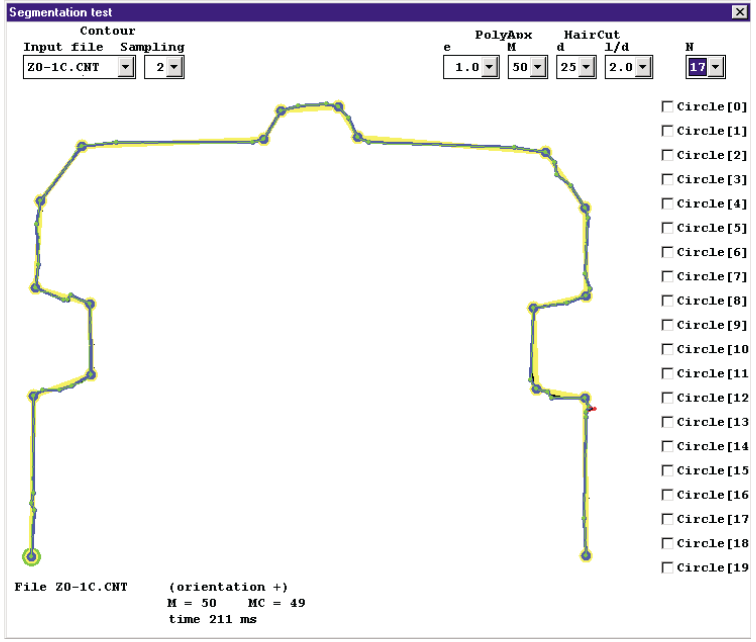
- Provides the back-light illumination where object shape appears as a silhouette.
- The principal is very often used in manufacturing for gauging and verifying correctness of a shape.
- The picture shows the illumination source (the smaller box containing LED diode and lens) and camera (a bigger box with cable, there is mirror inside). *Courtesy: Neovision s.r.o.*



Back-illuminated detail

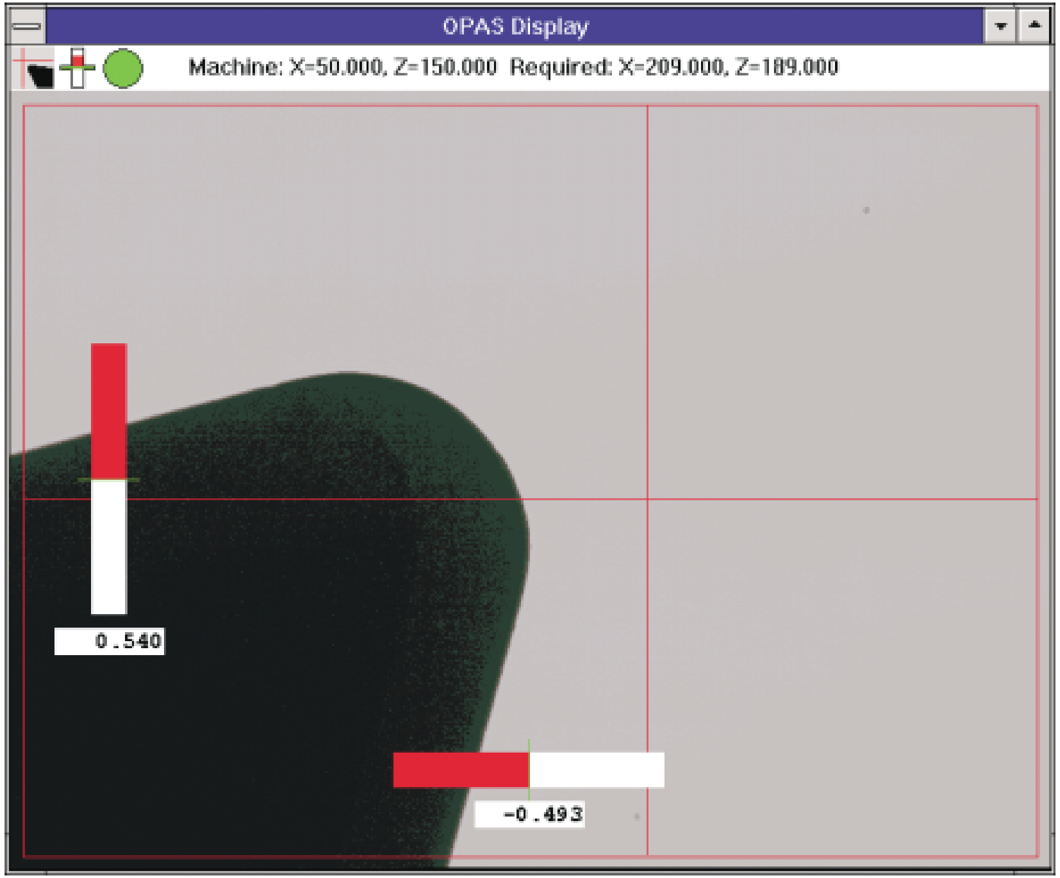
Example from the lathe-turning: *Courtesy: Neovision s.r.o.*

- Left image – a the back-illuminated detail showing imperfection and dust.
- Right image – automatically approximated shape allowing for automatic gauge check against a technical drawing.



Profile projector – User's screen

Another example: Automatic gauging of the lathe-turning knife from a sintered carbide.



Observations (or features) and segmentation

- There is an important question: Which observations (features in the pattern recognition terminology) distinguish regions corresponding to different objects in the image?
- Examples of features: intensity, color, shape of the region, texture, motion in video, disparity in stereo imaging.
- Primary features are provided by the sensor directly (e.g., intensity, color, depth in the range camera, temperature in far infrared (thermal) camera).
- More complicated secondary features have to be calculated from primary features as: texture parameters, shape parameters of the region, mutual relations between regions, motion parameters in video, stereo disparity, etc.

Use of a priori information

The more a priori information the better.

Illustrative examples:

- Required shape of the region.
- Required position, orientation.
- Known initial and final point of the boundary *(e.g., in the application analyzing shape of a polymer drop, where the polymer sample comes out a tube of known position).*
- Relation of the region considered to other regions with required properties *(e.g., above, inside).*

Examples from two application areas:

Remote sensing: Look for ships in the water. Typical properties of railway lines, highways (minimal curvatures). Rivers do not cross.

Medical: Blood vessels ar roughly parallel. Relate to anatomic atlas (model-based approach).

Major approaches to segmentation

- Threshold-based, according to a global property, usually intensity, where the global knowledge is represented by the intensity histogram.
- Spatial coherence-based ($\approx$ clustering of ‘tokens’).
  - Connecting, e.g., edgels because edges bear often an important information about objects (cf. human visual system).
  - Region merging/splitting. Regions come from aggregating pixels with similar properties (homogeneity criterion)
- Template matching – detection and fitting tokens in the image to a priori known template.
  - Parametric model detection, e.g., straight line, circle, ellipsis, . . .
- Unusual phenomena-based.
  - Camouflage detection based on an unusual texture.
  - Region segmentation for the image compression.

Thresholding

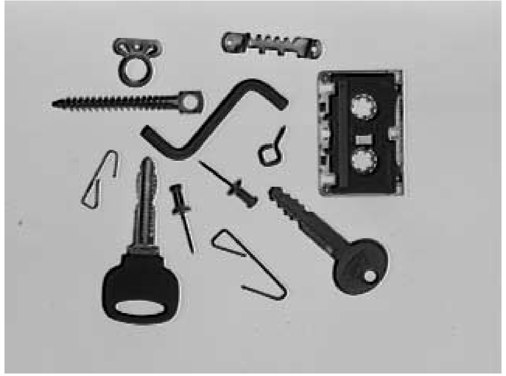
Input image $f(i,j)$, output image $g(i,j)$.

For each pixel (i,j)

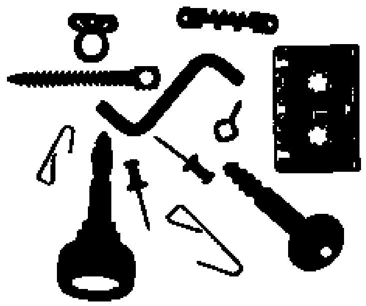
$$g(i,j) = \begin{cases} 1 & \text{for} \quad f(i,j) \geq \mathrm{Threshold}\,, \\ 0 & \text{for} \quad f(i,j) < \mathrm{Threshold}\,. \end{cases}$$

+ Simple technique, long time and more often used.
+ Easy in hardware, intrinsically parallel.
- The threshold is a parameter which is difficult to adjust automatically in general.
- Works only for subclass of images in which objects are distinct from background in intensity.

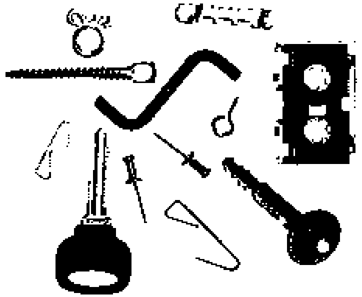
Example, threshold influence



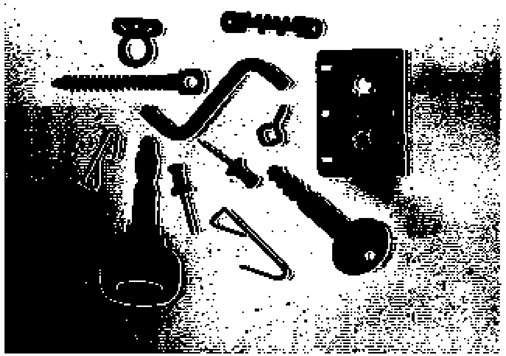
Original image.



Threshold segmentation.



Threshold too low.



Threshold too high.

Thresholding, modifications

Local adaptive thresholds, e.g. divide an image into subimages and find threshold in each of them.

Band thresholding, let D be a set of intensities, e.g. an interval of intensities

$$g(i,j) = \begin{cases} 1 & \text{for } f(i,j) \in D\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

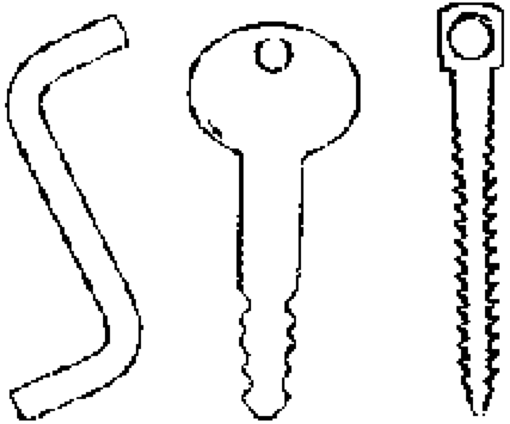
Multiple thresholds.

Semi-thresholding, to suppress background, useful in cases in which image is analysed by a human

$$g(i,j) = \begin{cases} f(i,j) & \text{for} \quad f(i,j) \geq \mathrm{Threshold}\,, \\ 0 & \text{for} \quad f(i,j) < \mathrm{Threshold}\,. \end{cases}$$

Example. Border regions by the band thresholding

Original image.



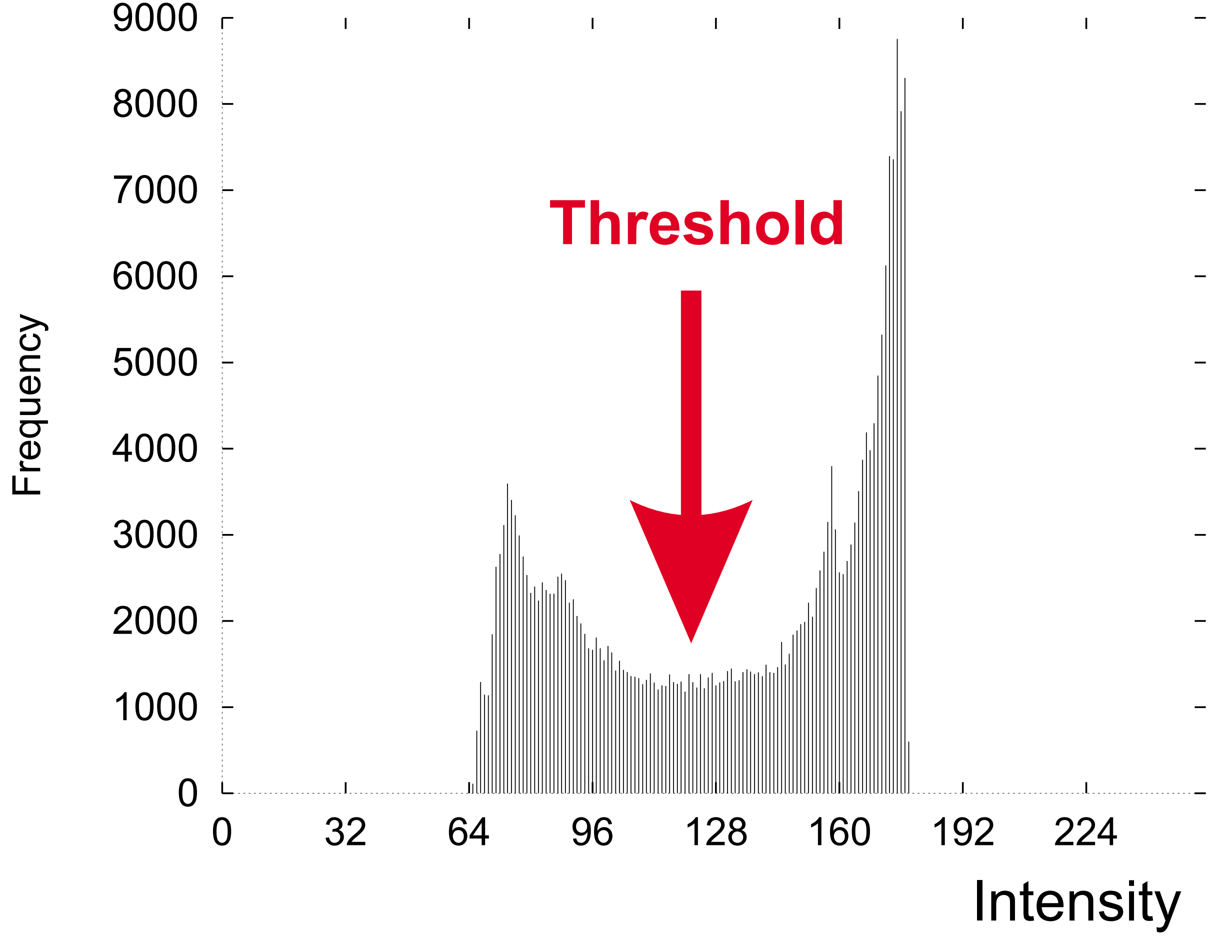
Border regions detected.

Automatic threshold detection, Use of histogram

p-tile thresholding, if we know that the objects cover $1/p$ of the image, e.g. printed characters on a sheet $\Longrightarrow 1/p$ area of a histogram.

Histogram shape analysis, distinct objects on background correspond to a bi-modal histogram. Find middle between the modes.

Automatically found threshold according to a bi-modal histogram



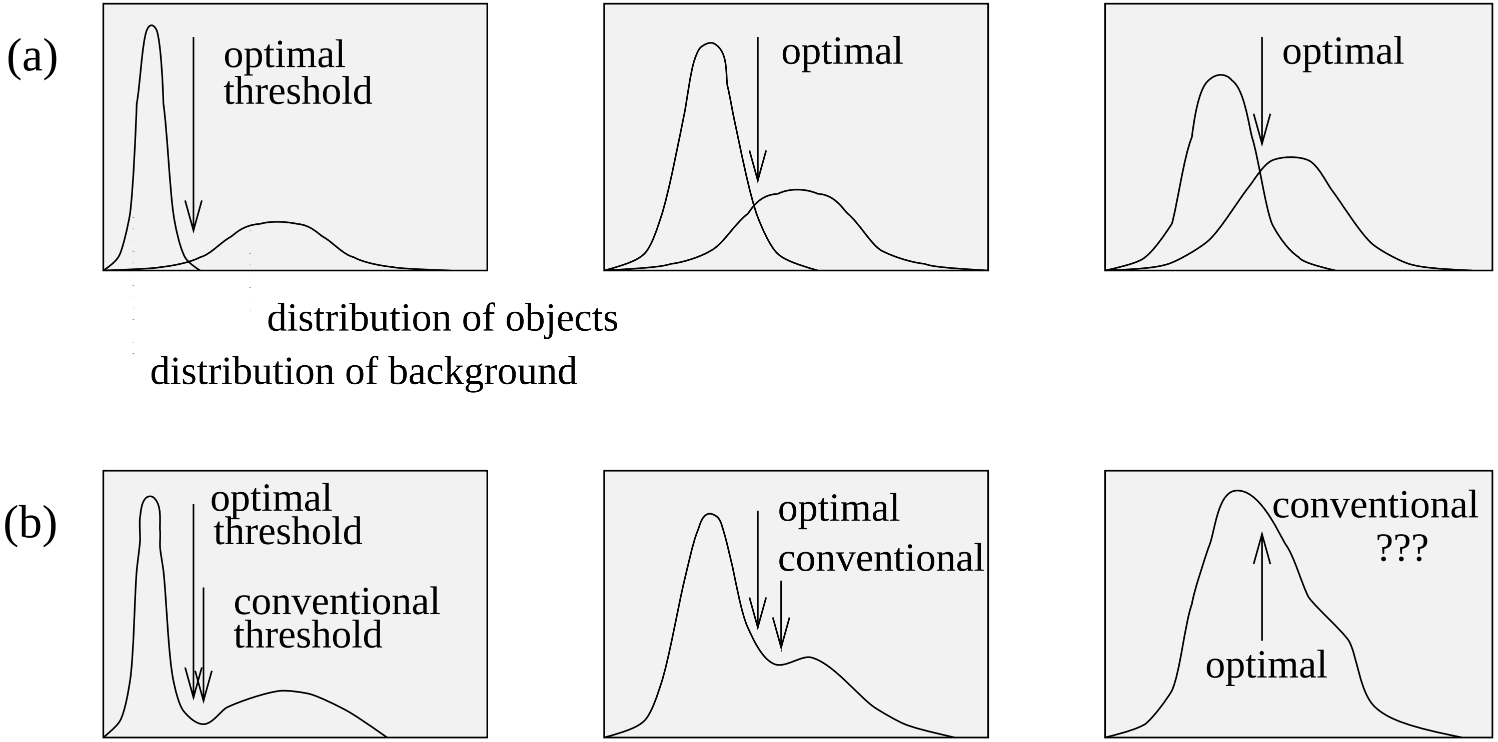
Smoothing of the ragged histogram

- Let $h_f(i)$ be a histogram which is ragged. Intensities $i = 0, \ldots, i_{\max}$.
- Smoothing before finding the threshold automatically helps to avoid the problem with many local minima.
- The 1D smoothing using sliding averaging of the window size $2K + 1$ is often used.
- The new histogram $h'_f(i)$ is calculated as

$$h'_f(i) = \frac{1}{2K+1} \sum_{j=-K}^{K} h_f(i+j)\,, \quad i = K, \ldots, i_{\max} - K\,.$$

Optimal thresholding by a mixture of Gaussians

Motivation:



Local thresholds by mixture of Gaussians

- h_{region} – local histogram.
- h_{model} – approximation of a histogram by n Gaussian distributions,

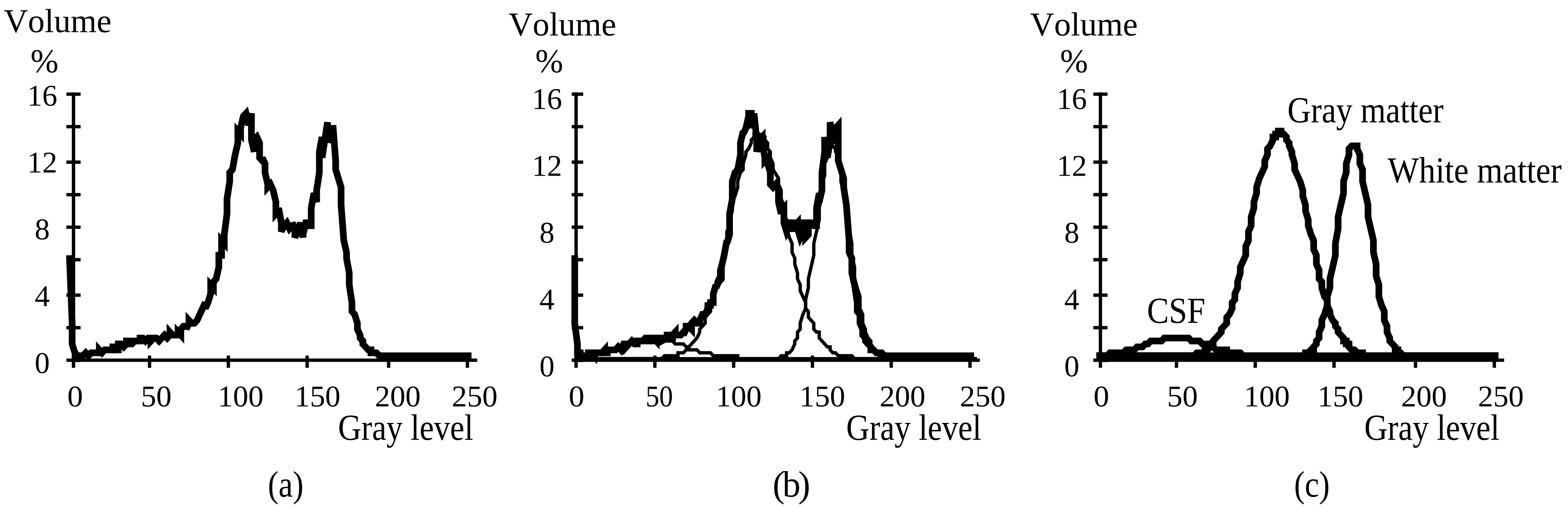
$$h_{\text{model}}(g) = \sum_{i=1}^{n} a_i e^{\left(-\frac{(g-\mu_i)^2}{2\sigma_i^2}\right)} .$$

- The minimized optimization criterion F,

$$F = \sum_{g \in G} \Big(h_{\text{model}}(g) - h_{\text{region}}(g) \Big)^2 .$$

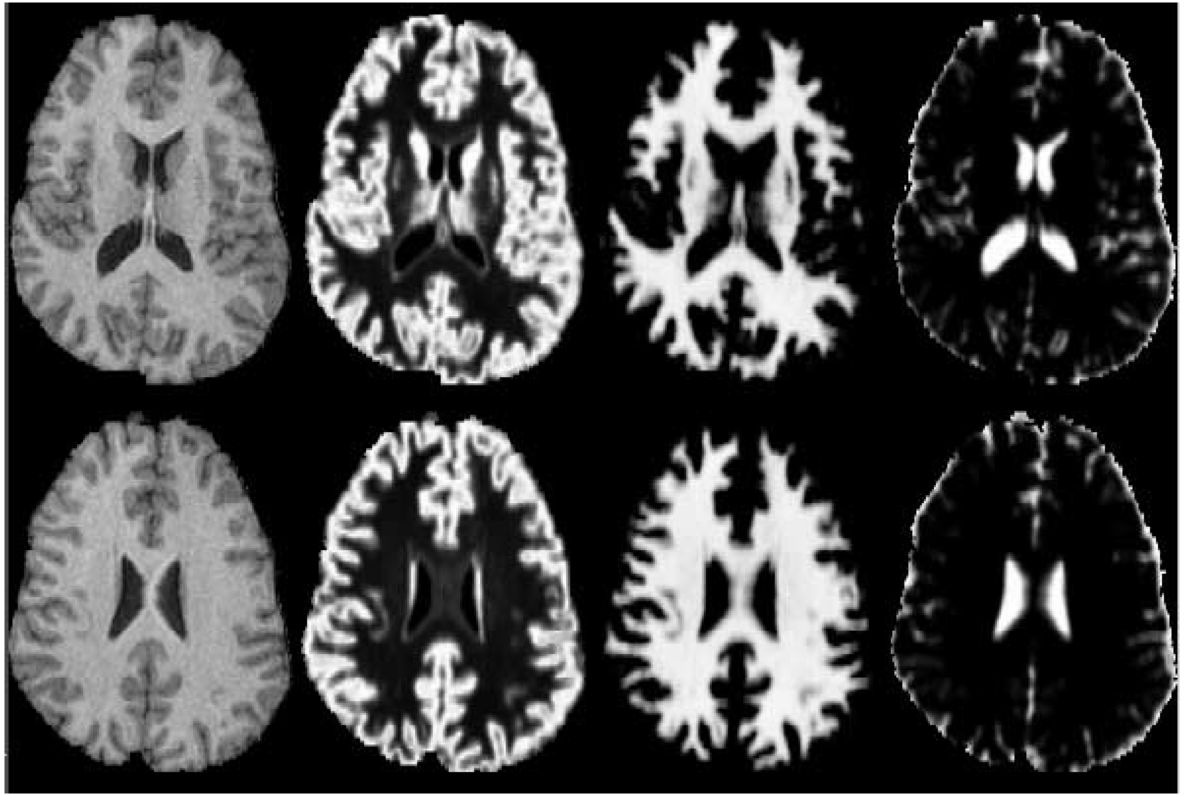
Example, Segmentation of the brain MR

- Input: T1-weighted NMR images.
- Desired classes: white matter, grey matter, celebro-spinal fluid (CSF)



Courtesy: Milan Šonka, University of Iowa.

Brain MR, Segmentation result



original gray matter white matter CSF