Wrocław University of Technology

An Image Generalization Technique - The Key to Finding Similar Images

Wojciech Tarnawski, Chair of Systems and Computer Networks, Faculty of Electronics

Łukasz Mirosław, Roman Pawlikowski, Krzysztof Ociepa,

Institute of Informatics

Outline

- **Introduction**
- **How do we interpret image-to-image similarity ?** – „enigmas of perception” -> Gestalt Theory->Multi-Scale Approach
- **GENeralization-oriented image SEGmentation (GEN-SEG)** - compromise between interpretation of all image details and the interpretation where certain details could be omitted
- The functional flowchart of the proposed **image retrieval system** composed of GEN-SEG algorithm and MPEG-7 visual descriptors
- **Results**

How do we interpret image-to-image similarity ?

…why do we interpret stimuli arriving at our retina as mountains, ocean, sky, car or cat ?

The answer is a little irritating: What is more natural than recognizing „mountain” in mountaing image or „ocean/sky/car/cat” in ocean+sky/car+cat image …

How groups of pixels are built into spatially extended visual object is one of the major **„enigmas of perception”…**

From Image Perception to Gestalt Theory

… to the best of our knowledge : visual reconstruction is governed by the **Gestalt Theory**(1) related with two kind of laws(2):

1) Grouping laws – *starting from the atomic local level recursively construct larger groups in the perceived image. Each grouping law focuses on a single quality (color, direction, shape, texture…)*

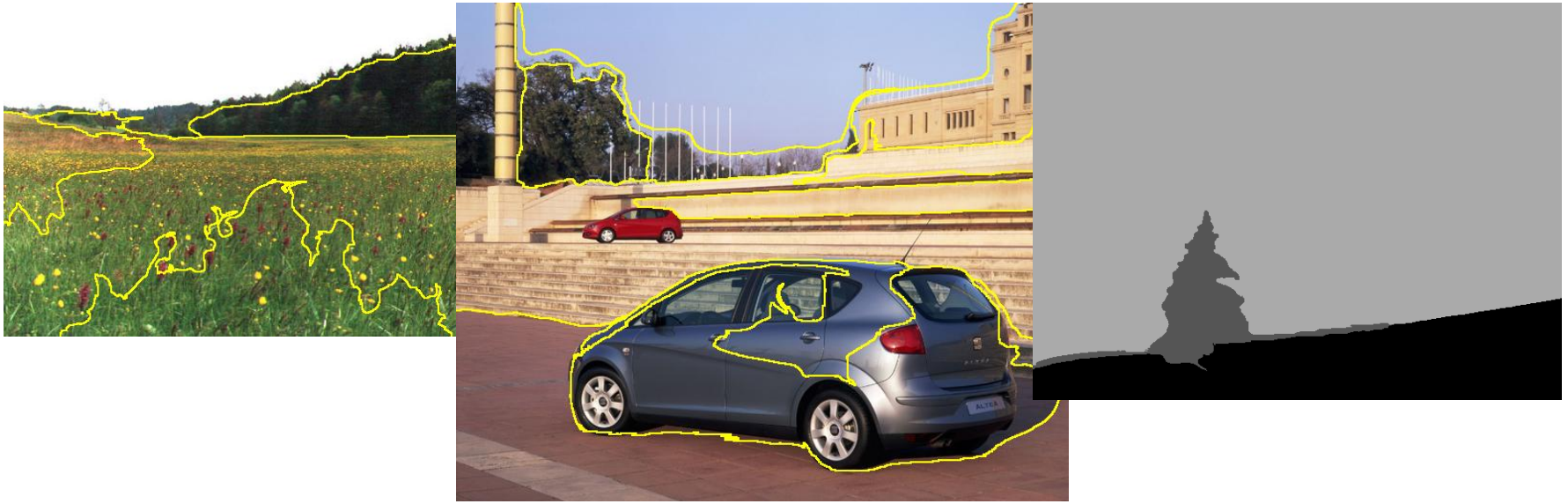
2) Collaboration and conflicts – *laws governing the interaction of grouping laws*

(1)„Gestalt" (German word) - "essence or shape of an entity's complete form"

(2) A. Desolneux, L. Moisen, J.M Morel – „*From Gestalt Theory to Image Analysis A Probabilistic Approach"* , Interdisciplinary Applied Mathemathics , vol. 34, Springer

From Gestalt Theory to Multi-Scale approach

In image understanding/retrieval it is difficult to find a compromise between interpretation of all image details and the interpretation where certain details could be omitted.

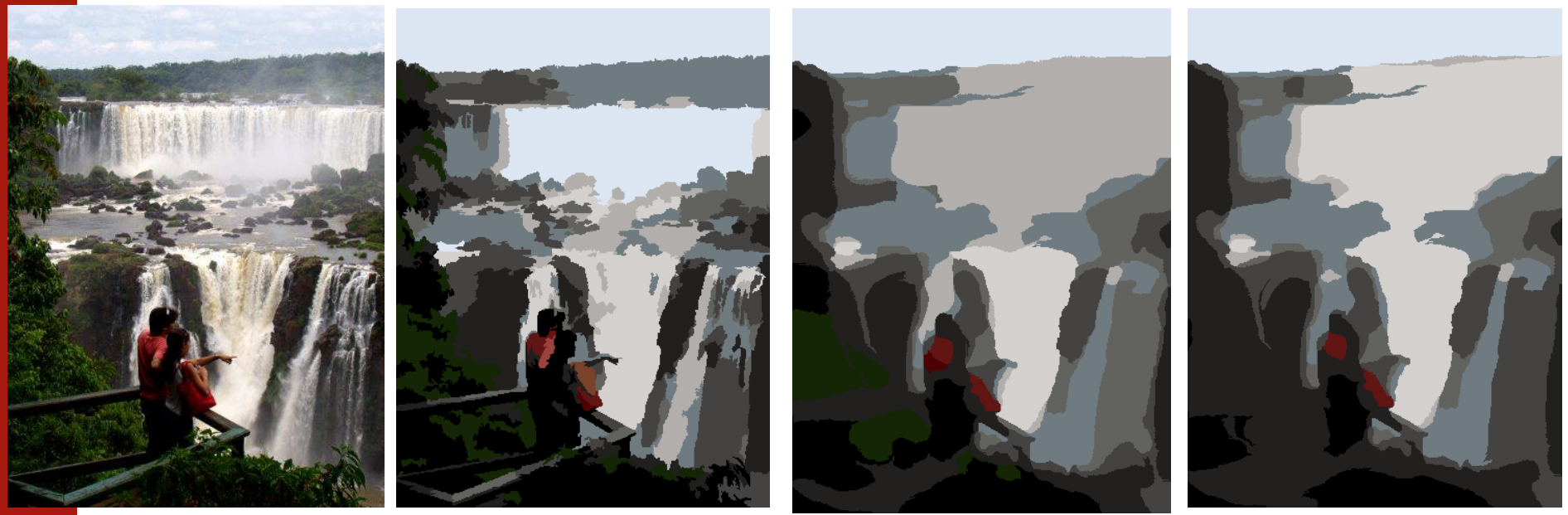


Conclusion: Without apriori knowledge there is no automated way to determine which details can be disregarded and which objects are large enough to be treated as significant.

From Gestalt Theory to Multi-Scale approach

Therefore, a **SCALE** should be considered as a parameter that changes dynamically and generates images at different scales. With changing scale the degree of precision also changes. Generated images form a so-called

Multi-Scale Image Representation



Multi-Scale Approach : Non-linear isotropic diffusion

Here, a number of images are generated in the process of convolving the original image with the Gaussian kernel with the *t*-variance (scale):

$$I(x, y, t) = I_0(x, y) \otimes G(x, y, t)$$

The set of derivative images *I(x,y,t)* is equivalent to concurrent solutions of the heat transport problem or diffusion on the plane.

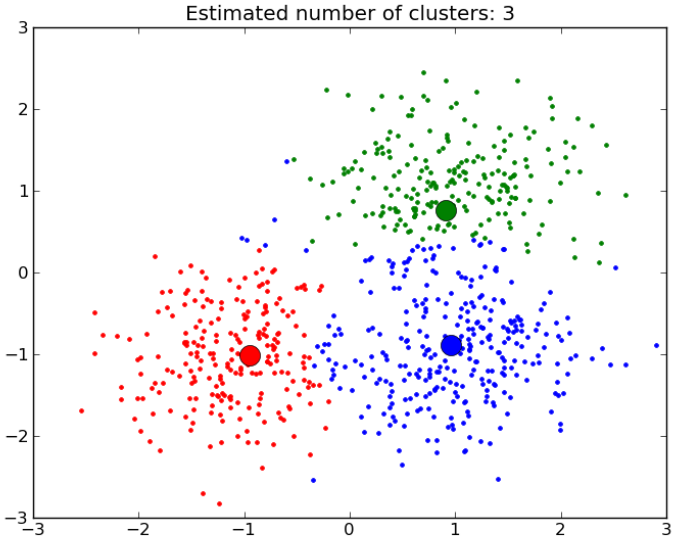
$$I_t = \nabla^2 I = I_{xx} + I_{yy}$$

Since, the convolution operation smoothes region boundaries we decided to use edge-preserving isotropic diffusion:

$$I_t = div(D(x,y,t) \cdot \nabla I) = D(x,y,t)\nabla^2 I + \nabla D \cdot \nabla I$$

Mean-shift segmentation

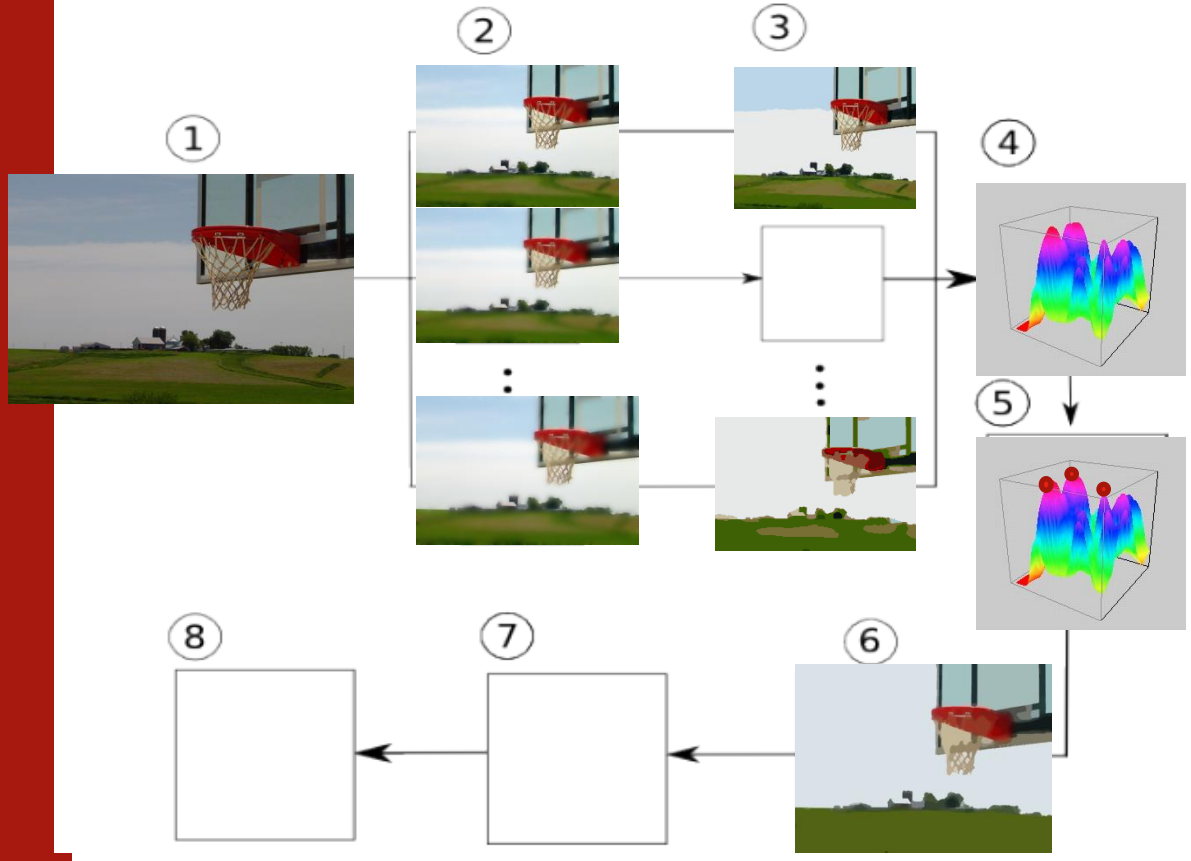
The ”mean-shift”-based image segmentation algorithm(3) is a non-parametric clustering in 5D image space (3D color space + 2D planar space). The method does not require to know the number and the shape of clusters.



(3) D. Comaniciu and P. Meer, *“Mean shift: A robust approach toward feature space analysis”* IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603–619, 2002.

GENeralization-oriented image SEGmentation (GEN-SEG) – key-solution to finding similar images

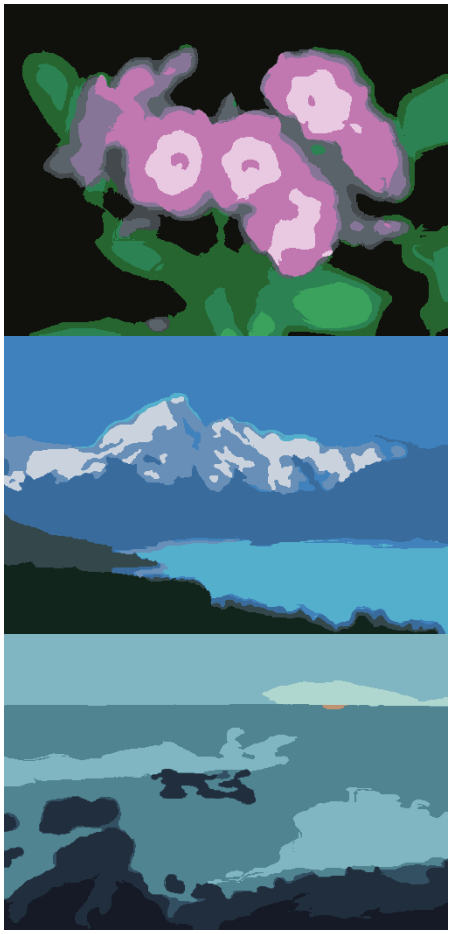
The aim of the proposed segmentation is to partition the image into non-overlapping most – meaningful image-regions to receive a very generalized (coarse) view of natural images.



1) Input image
2) Isotropic diffusion step
3) Mean-shift based segmentation
4) **Accumulating the clustering results over all scales**
5) Mode based clustering of values collected in accumulator space
6) „Visualization” of the clusters detected in previous step - mapping input pixels into set of cluster-labels
7) **Region merging taking into account region size and planar co-occurency**
8) Output image

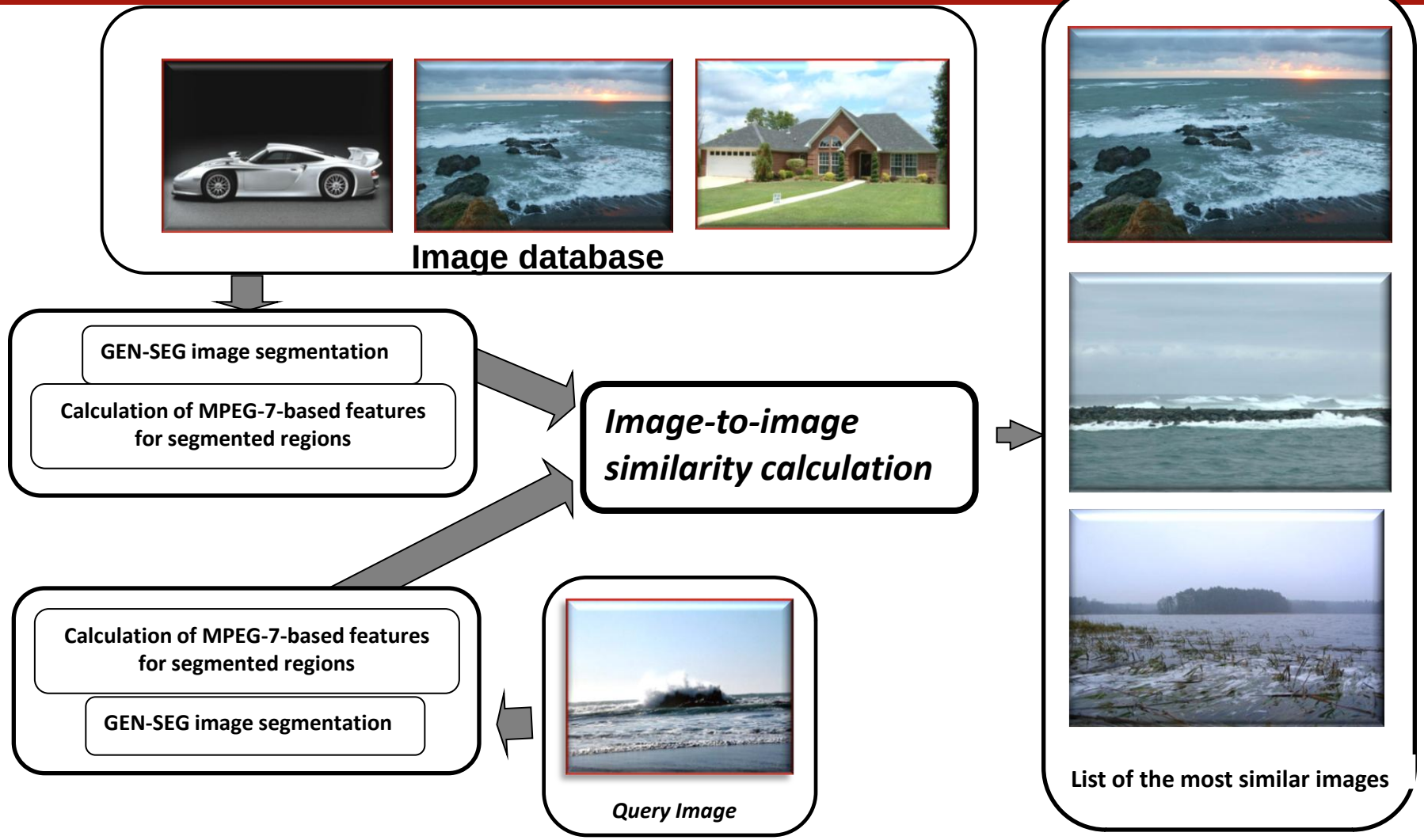
GEN-SEG segmentation vs Gestalt Theory

„Visualization” of clusters detected in the accumulator space : **Grouping laws**



The most-meaningful (merged) regions: **Collaboration laws**

The functional flowchart of image retrieval system



MPEG-7 visual descriptors

The Moving Picture Experts Group (MPEG) defined few visual descriptors and the appropriate distance measures (MPEG-7 standard):

Scalable Color Descriptor – color histogram in HSV color space that is encoded by a Haar transform.

Color Layout Descriptor – represents the spatial distribution of the color of visual signals.

Edge Histogram – represents the spatial distribution of five types of edges, namely four directional edges and one non-directional

Texture Browsing – 5D descriptor related to perceptual characterization of texture in terms of regularity , coarseness and directionality

Region Shape – describes the shape of an object in image. The descriptor is robust to noise that may be introduced in the process of segmentation

Definition of the image-to-image similarity

We have defined the measure of the similarity between images by composition of distance metrics defined in MPEG-7 standard. The procedure consists of the following steps:

1. The calculation of region-to-region distances between the query image and images in the database for chosen set of MPEG-7 descriptors

2. Sort the list of distances for all regions from query-image

   Democracy: each region belongs to certain image in the database and votes for it

3. Accumulate the votes and pick the best candidates

Results

We have done the preliminary tests in database MGV-2006 that includes about 70 annotations

We compared the proposed image retrieval scheme with other image retrieval algorithms designed and developed in group of prof. Kwasnicka using F-score metric.

Our method gives the highest values of F-score for 9 annotations in MGV-2006 database.

Thank You

Wojciech Tarnawski, Ph.D.
Wroclaw University of Technology
wojciech.tarnawski@pwr.wroc.pl