# A DATA-DRIVEN REGION DETECTOR FOR STRUCTURED IMAGE SCENES

*Elena Ranguelova*

Netherlands eScience Center
Amsterdam, The Netherlands

## ABSTRACT

## 1. INTRODUCTION

## 2. DATA-DRIVEN MORPHOLOGY SALIENT REGIONS DETECTION

MSER and MSSR decompose a gray-scale image into binary cross-sections and evaluate the stability of the connected components (CCs) or accumulate saliency masks. DMSR starts with a data-driven binarization, producing one binary image and thus transforming the problem into a binary saliency.

### 2.1. Binary Salient Regions Detection

We claim that the perceptual saliency in a binary image of a structured scene $\mathbf{B} : \mathcal{D} \subset \mathcal{Z}^2 \rightarrow \{0,1\}$ (1-white, 0-black) is only due to the spatial layout of the image regions [1]. We have defined four types of salient regions called *holes, islands, indentations* and *protrusions*. The two types of *inner salient structures (ISS)* are *holes* – set of connected black pixels entirely surrounded by white pixels, and *islands* – set of connected white pixels surrounded by black ones, i.e. inverse of holes. A significant connected component $\mathcal{B}^1$ is defined as a CC with area proportional to the image area by $\Lambda$. The radius of the morphological structuring element is $r$ and the area opening parameter for noise filtering is $\lambda$. In addition, the two *boundary salient structures (BSS)* are *protrusions*– set of white pixels on the border of a significant CC, which if pinched off from the CC, will increase its boundary with no more than $2\pi r$, and the *indentations*– inverse protrusions.

These four types are also valid for the MSSR detector. The regions are obtained from the binary image $\mathbf{B}$ by morphological operations: hole filling, top hat and area opening, for details see [2, 1]. The ISS are similar to the definition of the MSER+ and MSER- regions [3]. In this paper, detectors using only ISS, i.e., directly comparable to MSER, are denoted by DMSR/MSSR, while DMSRA/MSSRA are detectors using all region types. These definitions are summarized in Table 1 and the exact shaped regions from a synthetic $100 \times 100$ binary image with parameters $\Lambda = 100$, $r = 5$ and $\lambda = 10$ are shown on Fig. 1.

| | |
|---|---|
| **ISS** | A CC $S^i_{fb} = \{\mathbf{p} \in \mathcal{D}, \forall \mathbf{p} = foreground,$ $\forall \mathbf{q} \in \partial S^i_{fb}, \mathbf{q} = background, \mathbf{q} \notin \partial \mathbf{B}\}$, |
| 2 types | $S^i_{10}$ (islands), $S^i_{01}$ (holes); $\mathbf{S}^i = S^i_{01} \cup S^i_{10}$ |
| **BSS** | $S^b_{fb} : \{\mathbf{p} \in S^b_{fb} \subset \mathcal{B}^f, \forall \mathbf{p} = foreground,$ $\mathbf{q} \in \partial S^b_{fb} \subset \partial \mathcal{B}^f, \forall \mathbf{q} = background\}$, $\lvert\partial \mathcal{B}^f\rvert - \lvert\partial(\mathcal{B}^f \backslash S^b_{fb})\rvert < 2\pi r$ |
| 2 types | $S^b_{10}$ (protr.), $S^b_{01}$ (indent.); $\mathbf{S}^b = S^b_{01} \cup S^b_{10}$ |
| **Regions** | $\mathbf{S} = \mathbf{S}^i$ (DMSR); $\mathbf{S} = \mathbf{S}^i \cup \mathbf{S}^b$ (DMSRA) |

**Table 1**. Binary saliency definitions used in Section 2.1.

**Fig. 1**. Binary salient regions detection. Color coding: holes- blue, islands- yellow, indentations - green, protrusions- red.

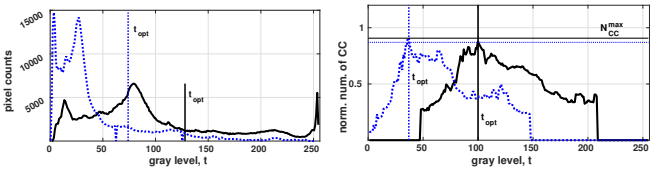
### 2.2. Data-driven binarization

Any gray-scale image $\mathbf{I} : \mathcal{D} \subset \mathcal{Z}^2 \rightarrow \mathcal{T}$, where $\mathcal{T} = \{0, 1, \ldots, t_{max}\}$ and $t_{max} = 2^n - 1 = 255$ is the maximum gray value encoded by $n = 8$ bits, can be decomposed into cross-sections at every possible level $t$: $\mathbf{I} = \sum_{t \in \mathcal{T}} CS_t(\mathbf{I})$. Obtaining a section at level $t$ is equivalent to thresholding the image at threshold $t$: $CS_t(\mathbf{I}) = 1.(\mathbf{I} > t) + 0.(\mathbf{I} < t)$ is a binary image. Three sets of connected components in $CS_t(\mathbf{I})$ are defined: $\mathcal{A}_t$- *all*, $\mathcal{L}_t$- the *large* and $\mathcal{V}_t$- the *very large* CCs. The size of each CC category is defined by $\Lambda_{\mathcal{L}}$ and $\Lambda_{\mathcal{V}}$ fraction of the image area $A_{\mathbf{I}}$. Let us denote the normalized number of elements in a set ($|\cdot|$) by $\|\cdot\| = |\cdot| / \max_{t \in \mathcal{T}} |\cdot|$. Finding the optimal threshold $t_{opt}$ is then defined as:

$$t_{opt} = \arg\max_{t \in \mathcal{T}} (w^{\mathcal{A}} \|\mathcal{A}_t\| + w^{\mathcal{L}} \|\mathcal{L}_t\| + w^{\mathcal{V}} \|\mathcal{V}_t\|),$$

where $w^{\cdot}$ are the weights per set of CCs.

The standard Otsu thresholding does not select a single stable $CS$, while choosing $t_{opt}$ ensures stable number of regions across transformations (see Figs. 2 and 3 for lighting).




(a) Otsu (b) Max number CC

**Fig. 2**. Finding the optimal threshold for two images from the 'Leuven' sequence (Oxford dataset, lighting): the base image- solid black line, the forth image - dotted blue line.

(a) (b) (c)

**Fig. 3**. Binarization of two images of the 'Leuven' sequence (lighting). Top row- base image, bottom row- forth image; (a) gray scale; (b) Otsu binarization, (c) proposed binarization.

After the data-driven binarization, the DMSR detector finds the set of affine-covariant regions $\mathbf{S}$ from the single binary image $CS_{t_{opt}}$ as was described in Section 2.1 and [2, 1]. DMSR produces fewer non-overlapping and perceptually salient regions (visualized by their equivalent ellipses, not exact shapes) compared to MSER (see Figures 4 and 5).

## 3. PERFORMANCE EVALUATION

*performance evaluation here...*

**Fig. 4**. Region detectors on the base image of the 'Graffiti' sequence, Oxford dataset. Left: MSER, right: DMSR

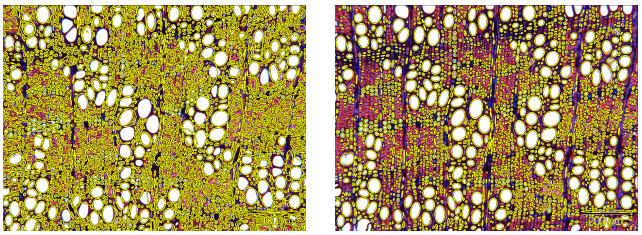


**Fig. 5**. Salient region detectors on microscopy wood images. Left: MSER (every second region is shown), right: DMSR

## 4. REFERENCES

[1] E. B. Ranguelova and E. J. Pauwels, "Saliency Detection and Matching for Photo-Identification of Humpback Whales," *International Journal on Graphics, Vision and Image Processing*, 2006.

[2] E. B. Ranguelova and E. J. Pauwels, "Morphology-Based Stable Salient Regions Detector," in *Proceedings of International Conference on Image and Vision Computing New Zealand*, 2006, pp. 97 – 102.

[3] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust Wide Baseline Stereo from Maximally Stable Extremal Regions," in *Proceedings BMVC*, 2002, pp. 36.1–36.10.