

DATA-DRIVEN REGION DETECTOR FOR STRUCTURED IMAGE SCENES

Elena Rangelova

Netherlands eScience Center
Amsterdam, The Netherlands

ABSTRACT

A Data-driven Morphology Salient Regions (DMRS) detector, related to our MSSR detector is proposed. It demonstrates comparable repeatability to the best-known MSER detector on standard structured scenes and better resolution invariance on a high-resolution benchmark. This is achieved via significantly smaller number of detected regions- a much desired property in the big data era. A data-driven binarization algorithm gives compact image representation, subsequently analyzed for saliency using morphology. Also, a new benchmark dataset, 'OxFrei', for transformation-independent detection evaluation is introduced. While MSER is an excellent detector for generic applications, the DMSR is geared towards the analysis of scientific imagery for detecting precisely semantically meaningful regions. This is a key property in emerging domains such as animal and plant biometrics, where computer vision becoming the vital technology to aid the wild-life preservation efforts. In this paper, DMSR is demonstrated to better detect identifying structures in marine animals and wood microscopy images.

Index Terms— region detection, data-driven, morphology, structured scenes, scientific visual analytics

1. INTRODUCTION

Finding reliably and repeatedly correspondences between two images of the same object or scene, taken under different viewpoints and acquisition conditions, is the first fundamental step in numerous computer vision applications: wide baseline stereo matching, image retrieval, model-based recognition, visual mining, object categorization, etc. An important class of features are the stable and salient regions, which correspond to the same image patches for different viewpoints, detected independently for each image. The region detectors must be *covariant* (or often called *invariant* in the literature) to a class of transformations, usually *affinity* and various photo-metric distortions.

While the most research has been focused on the generic applications, the emerging fields of *animal and plant biometrics*, is attracting more attention of the community [1, 2]. It becomes clear that computer vision is the vital technology enabling the work of scientists, ecologists, marine biologists,

etc in the big data era. An important question for these domain scientists, along with the individual or species photo-identification, is to obtain automatic reliable measurements of semantically meaningful regions, extracted from usually highly structured images. The state-of-the-art generic region detectors often do not satisfy this requirement and research on new methods is needed.

The performance of the region detectors is measured on evaluation benchmarks, [3, 4, 5]. Despite the growing need, there is a shortage of publicly available benchmarks, which enable measuring the robustness to image transformations independently of the image content.

In this paper, we tackle the above identified problems by developing an affine-covariant interest regions detector for structured images, the *Data-driven Morphology Salient Regions (DMRS)*. It has properties making it suitable for enhancing the scientific imagery analytics in the big data era (Figs. 1, 2). A new dataset, 'OxFrei', for performance evaluation is also introduced.

1.1. Related work

A decade ago, a performance evaluation paper by the Visual Geometry Group in Oxford, compared the existing affine-covariant region detectors, [3]. The authors created a reference dataset, which has become the standard evaluation benchmark since, the *Oxford dataset*. A clear conclusion of the comparative study was that the *Maximally Stable Extremal Regions (MSER)* detector is the best performing detector for *structured* scenes for many of the tested transformations, [6]. The MSER detector has become the de-facto standard in the field, for example is now an integral part of the MATLAB CVS Toolbox. Despite its success, the MSER detector has several drawbacks: it is sensitive to image blur; produces many nested regions (not coinciding with the perceptual saliency); the number of image regions is often large (undesirable for matching in large datasets) and the performance degrades with up to 25% with increase of the image resolution, [5]. Analysis of the MSER features in the geometric scale-space showed that the formulation of the stability criterion makes MSER biased towards regular shapes, which is a limitation when processing most natural images, [7].

Many researchers have proposed improvements to MSER,

though none of them increased the performance drastically. An MSER color extension, *Maximally Stable Color Region* is proposed in [8]. It outperforms a simple color MSER extension and a color blob detector on the Oxford dataset. Another enhancement of MSER with the Canny edge detector is introduced in [9]. The dilation operator is used on the detected edges to remove the ambiguous ones, making the interest regions more representative. The improved detector shows better performance than MSER for image classification in a bag of words framework, though the benchmark evaluation of repeatability is not reported. The MSER detector has been also extended to 3D, *Maximally Stable Volumes*, and used to successfully segment 3D medical images and paper fiber networks, [10].

In the context of humpback whale identification, we have developed the *Morphology based Stable Salient Regions* (MSSR) detector, [11, 12]. While achieving less on the repeatably scores on the Oxford dataset, it produces less number of regions (which is important for the efficiency of the following matching step) and much better coincides with the human salient perception. DMRS is simpler and faster than MSSR, while pertains the perceptually salient property, in contrast of the often redundant concentric MSER regions, see Fig.1. Figure 2 illustrates that MSER does not even detect useful regions on one of the leatherback images, while DMSR detects meaningful regions repeatedly.

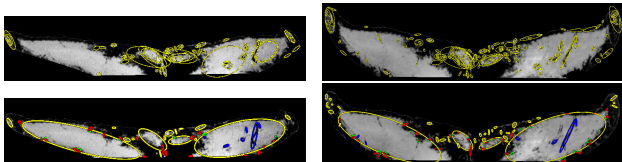


Fig. 1. Region detectors on two images of the tail of the same humpback whale. Top row: MSER, bottom row: DMSR.

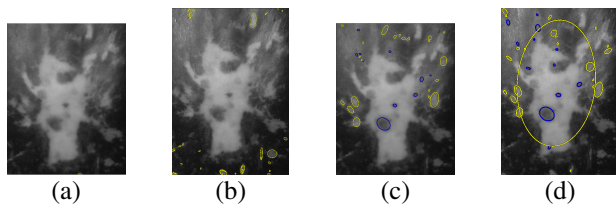


Fig. 2. Region detectors on two images of the pineal spot of the same leatherback turtle. (a),(b): MSER, (c),(d): DMSR.

There are not many evaluation benchmark datasets apart from the Oxford dataset, which is very small for nowadays standards: only 6 test sequences and 48 relatively low resolution images with known homographies and evaluation protocol, [3]. The *Freiburg dataset* contains 16 base images, which give 416 transformed higher resolution images, [4]. Unlike the naturally obtained images in the Oxford set, the Freiburg data have been generated by applying the transformations to

each of the base images. This is done to de-tangle the transformations from the image content. Its main drawback is the lack of proper documentation of the transformation parameters. The *TNT dataset* contains several versions of few image transformed sequences with increasing resolution from 1.5 MB to 8 MB, along with very highly accurate homographies, compatible with the standard evaluation protocol, [5]. The TNT set contains only viewpoint transformation, focusing on evaluation robustness to resolution.

1.2. Contributions

This paper makes the following contributions:

- Binarization algorithm, robust to image lighting and blur.
- DMSR salient regions detector for structure scenes. It is simpler and faster compared to MSSR, while the regions are also perceptually salient.
- DMSR gives much smaller and more stable number of regions compared to MSER, while having comparable or higher repeatability than MSER (for lighting, blur and increased resolution).
- More informative way of performance evaluation reporting in 3D for better insight in transformation covariance and higher resolution invariance.
- New open-source dataset 'OxFrei' combining good features of the Oxford and Freiburg datasets. The detector software is also open source.
- Demonstrated the potential of DMSR for scientific imagery analytics, mostly in the animal and plant biometric fields.

2. BINARY SALIENT REGIONS DETECTION

2.1. Binarization algorithm

3. DATA-DRIVEN MORPHOLOGY SALIENT REGIONS

4. PERFORMANCE EVALUATION

4.1. Standard datasets

4.1.1. Oxford dataset

4.1.2. Freiburg dataset

4.1.3. Combined dataset

4.1.4. TNT hi-res benchmark

5. CONCLUSIONS

6. REFERENCES

7. REFERENCES

- [1] H. Kuehl and T. Burghardt, "Animal Biometrics: quantifying and detecting phenotypic appearance," *Trends in*

Ecology & Evolution, vol. 28, pp. 432–441, 2013.

- [2] N. Kumar, P. Belhumeur, A. Biswas, D. Jacobs, W. Kress, I. Lopez, and J. Soares, “Leafsnap: A Computer Vision System for Automatic Plant Species Identification,” in *The 12th European Conference on Computer Vision (ECCV)*, October 2012, pp. 502–516.
- [3] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, “A comparison of affine region detectors,” *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43–72, November 2005.
- [4] P. Fischer, A. Dosovitskiy, and T. Brox, “Descriptor Matching with Convolutional Neural Networks: a Comparison to SIFT,” *CoRR*, vol. abs/1405.5769, 2014.
- [5] K. Cordes, B. Rosenhahn, and J. Ostermann, “High-Resolution Feature Evaluation Benchmark,” in *The 15th International Conference on Computer Analysis of Images and Patterns (CAIP)*, 2013, pp. 327–334.
- [6] 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.
- [7] R. Kimmel, C. Zhang, A. Bronstein, and M. Bronstein, “Are MSER Features Really Interesting?,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2316–2320, 2011.
- [8] Per-Erik Forssén, “Maximally Stable Colour Regions for Recognition and Matching,” in *CVPR*, 2007, pp. 1–8.
- [9] Sh. Wang, W. Wang, D. Liu, F. Gu, and B. Dickson, “Enhanced Maximally Stable Extremal Regions with Canny Detector and Application in Image Classification,” *Journal of Computational Information Systems*, vol. 10, no. 14, pp. 6093–6100, 2014.
- [10] M. Donoser and H. Bischof, “3D Segmentation by Maximally Stable Volumes (MSVs),” in *ICPR (1)*, 2006, pp. 63–66.
- [11] 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.
- [12] 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.