A DATA-DRIVEN REGION DETECTOR FOR STRUCTURED IMAGE SCENES

Elena Ranguelova

Netherlands eScience Center
Amsterdam, The Netherlands

ABSTRACT

Finding correspondences between two images of the same scene, taken from different viewpoints, with semantic features is a challenging problem. This paper proposes a Data-driven Morphology Salient Regions (DMSR) approach for detecting interest regions repeatedly. A binarization algorithm creates a compact image representation that is then analyzed for saliency using morphology. DMSR has comparable performance to the renowned Maximally Stable Extremal Regions (MSER) detector on structured scenes and better invariance to lighting, blur and on a high-resolution benchmark. This is achieved via significantly fewer detected regions, leading to better scalability. DMSR is shown to be a better choice than MSER for analysis of scientific imagery in the big data era, e.g., it detects precisely meaningful regions in images used for wild-life biometrics. The paper also introduces OxFrei, a dataset for transformation-independent detection evaluation.

1. INTRODUCTION

The first fundamental step in numerous computer vision applications (wide baseline stereo matching, image retrieval, visual mining, etc.) is to reliably and repeatedly find the correspondence between a pair of different images of the same scene [1, 2, 3]. One class of methods- *region detectors* find distinct (salient) regions, which correspond to the same image patches, detected independently in each image. The detectors must be *covariant* (often called *invariant*) to, usually, *affine* transformations and various photometric distortions.

A decade ago, a performance evaluation paper by the Visual Geometry Group in Oxford compared existing region detectors [4]. A clear conclusion of the comparison was that *Maximally Stable Extremal Regions (MSER)* is the best performing detector for *structured* scenes, e.g., those containing homogeneous regions with distinctive boundaries [1]. MSER has become the de-facto standard in the field, e.g., it is in the MATLAB Computer Vision Systems Toolbox and OpenCV. Despite its success, the detector has several drawbacks: it is sensitive to image blur; it produces nested and redundant regions and its performance degrades with the increase of image resolution [5]. Analysis in geometric scale-space have showed that the formulation of the region stability criterion makes MSER prefer regular shapes [6].

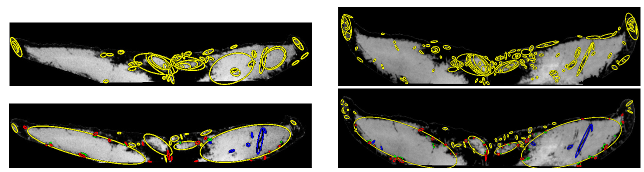
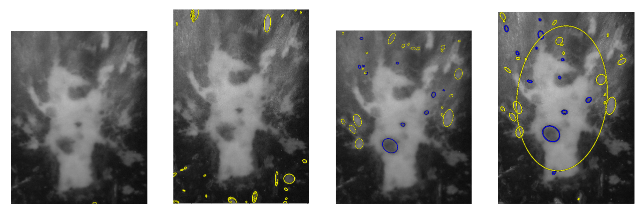


Fig. 1. Region detection on two images of the tail of the same humpback whale. Top row: MSER, bottom row: DMSR(All).



(a) (b) (c) (d)

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

Since then, many researchers have proposed improvements to MSER with no drastic increase of performance. An MSER color extension, *Maximally Stable Color Region*, outperforms both an MSER-per-color-channel combination and a color blob detector [7]. Improving the MSER region distinctiveness by morphological dilation on the detected Canny edges is proposed in [8]. The improved detector shows better performance in classification application, but evaluation of repeatability is not reported. MSER has been extended to *Maximally Stable Volumes* for successful segmentation of 3D medical images and paper fiber networks [9].

While most research has focused on generic applications, the emerging fields of *animal and plant biometrics* are attracting more attention [10, 11]. Computer vision is becoming a vital technology enabling the wild-life preservation efforts of ecologists in the big data era. Along with the individual or species photo-ID, the scientists wish to obtain reliable measurements of meaningful structures from images. The generic region detectors do not satisfy this need: Fig. 1, top row shows over-abundant or not semantic regions and Fig. 2, (a), (b) shows missed detection. On the contrary, our *Data-driven Morphology Salient Regions (DMSR) detector* finds the semantic regions (Fig. 1, bottom row and Fig. 2, (c), (d)).

Although crucial for the development of detectors, there is a shortage of evaluation benchmarks, especially for performance analysis independently of the image content. The standard *Oxford dataset* is very small: eight test sequences containing six (one base and five transformed) images of the same scene each. Every pair (base, transformed) is related via a given transformation matrix (homography) [4]. The *Freiburg dataset* contains 416 higher resolution images, generated by transforming 16 base images in order to de-tangle transformations from content [12]. The *TNT dataset* contains versions of the same viewpoint sequences with increasing resolution from 1.5 to 8 MPixel per image. Highly accurate image pair homographies are given. It is suitable for evaluating robustness to resolution rather than to transformations [5].

This paper contributes to solving the identified problems. We propose a new regions detector, DMSR, and made the software available as open source [13]. It is related to the *Morphology-based Stable Salient Regions (MSSR)* detector that we developed in the context of humpback whale identification [14, 15]. DMRS includes a binarization, robust-to-lighting-and-blur, that yields a much smaller number of regions and is more stable across transformations. It has similar or higher (lighting, blur and increased resolution) repeatability compared to MSER, while detecting non-redundant perceptually salient regions (Fig. 1). Also, we compose and share an openly available dataset, OxFrei, combining the natural homographies of the Oxford and the higher resolution images of the Freiburg datasets [13].

2. REFERENCES

[1] 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.

[2] Foncubierta-Rodrguez et al., "Region-based volumetric medical image retrieval," in *SPIE Medical Imaging: Advanced PACS-based Imaging Informatics and Therapeutic Applications*, 2013.

[3] S. Escalera, P. Radeva, and O. Pujol, "Complex salient regions for computer vision problems," in *CVPR*, 2007.

[4] K. Mikolajczyk et al., "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43–72, November 2005.

[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] R. Kimmel et al., "Are MSER Features Really Interesting?," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2316–2320, 2011.

[7] Per-Erik Forssén, "Maximally Stable Color Regions for Recognition and Matching," in *Computer Vision and Pattern Recognition (CVPR)*, 2007, pp. 1–8.

[8] Sh. Wang et al., "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.

[9] M. Donoser and H. Bischof, "3D Segmentation by Maximally Stable Volumes (MSVs).," in *ICPR (1)*, 2006, pp. 63–66.

[10] H. Kuehl and T. Burghardt, "Animal Biometrics: quantifying and detecting phenotypic appearance," *Trends in Ecology & Evolution*, vol. 28, pp. 432–441, 2013.

[11] N. *et al.*, Kumar, "Leafsnap: Computer Vision System for Automatic Plant Species Identification," in *The 12th European Conference on Computer Vision (ECCV)*, October 2012, pp. 502–516.

[12] P. Fischer, A. Dosovitskiy, and T. Brox, "Descriptor Matching with Convolutional Neural Networks: a Comparison to SIFT," *CoRR*, vol. abs/1405.5769, 2014.

[13] E. Ranguelova, "Large scale imaging: Data, software, results," DOI: http://dx.doi.org/10.5281/zenodo.45156, Jan. 2016.

[14] 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.

[15] 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.