

Are MSER Features Really Interesting?

Ron Kimmel, *Fellow, IEEE*, Cuiping Zhang,
Alexander M. Bronstein, *Sr. Member,*
IEEE, and Michael M. Bronstein,
Sr. Member, IEEE

Abstract—Detection and description of affine-invariant features is a cornerstone component in numerous computer vision applications. In this note, we analyze the notion of maximally stable extremal regions (MSERs) through the prism of the curvature scale space, and conclude that in its original definition, MSER prefers regular (round) regions. Arguing that interesting features in natural images usually have irregular shapes, we propose alternative definitions of MSER which are free of this bias, yet maintain their invariance properties.

Index Terms—MSER, feature detector, affine invariance, stable region, correspondence.

1 INTRODUCTION

In recent years, feature descriptors extracted through linear scale-space analysis of an image have proven to be a powerful tool in object matching and recognition [1]. One of the most popular descriptor is the *scale-invariant feature transform* (SIFT) introduced by Lowe [2]. It first locates points of interest in a linear scale space, and then assigns a descriptor vector constructed as local histograms of image gradient orientations around the point. The descriptor itself is oriented by the dominant gradient direction, which makes it rotation invariant. SIFT uses linear scale space in order to search for feature points that appear at multiple resolutions of the image, which makes the method also scale invariant.

One of the main disadvantages of SIFT is that it is not affine invariant (see a recent work of Yu and Morel [3] on an affine-invariant version of SIFT). Affine invariance is important in image analysis since a more general class of viewpoint transformations can be approximated as local affine transformations of the image. An affine-invariant alternative to the SIFT widely used in computer vision applications is the *maximally stable extremal region* (MSER) [4]. This approach extracts stable regions from the image by considering the change in area with respect to the change in intensity of a connected component defined by thresholding the image at a given gray level. The change of area, normalized by the area of the connected component, is used as the stability criterion. The area ratio is invariant to affine transformations and so does the extracted region after appropriate canonization.¹ Benchmarks comparing the MSER, SIFT, other approaches, and affine-invariant

1. See [5] and [6] for a closely related approach that also allows for the analysis of contour segments, as well as [7] and [8] for an axiomatic framework of differential affine-invariant signatures of planar shapes.

- R. Kimmel is with the Department of Computer Science, Technion-Israel Institute of Technology, Haifa 32000, Israel. E-mail: ron@cs.technion.ac.il.
- C. Zhang is with the CMART Systems Inc., 404 Saratoga Ave., Santa Clara, CA 95050. E-mail: cuiping.zhang@gmail.com.
- A.M. Bronstein is with the School of Electrical Engineering, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: bron@eng.tau.ac.il.
- M.M. Bronstein is with the Institute of Computational Science, Faculty of Informatics, Università della Svizzera Italiana, Lugano 6900, Switzerland. E-mail: michael.bronstein@usi.ch.

Manuscript received 8 Sept. 2009; revised 10 Oct. 2010; accepted 12 May 2011; published online 22 June 2011.

Recommended for acceptance by S. Belongie.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2009-09-0600.

Digital Object Identifier no. 10.1109/TPAMI.2011.133.

alternatives thereof [9], [10] show that SIFT performs well for planar objects (like a graffiti wall) while the MSER performs better in most scenarios involving less trivial objects.

In this paper, we relate MSER to geometric scale-space analysis and image evolution by the level set curvature flow. We observe that the stability criterion in the original formulation of MSER prefers regular regions and, arguing that interesting features in natural images usually have irregular shapes, propose alternative definitions of MSER which are free of this bias, yet maintain their invariance properties. The rest of this paper is organized as follows: In Section 2, we briefly overview the theory of image representation as level sets and curve evolution. In Section 3, we formulate the MSER feature detection and discuss feature normalization algorithms. In Section 4, we discuss the drawbacks of the MSER stability criteria that make it sensitive to blur and tend to prefer round regions, and define alternative stability criteria. Section 5 shows experimental results. Finally, Section 6 concludes the paper.

2 IMAGE AS A COLLECTION OF LEVEL SETS

Let $X \subset \mathbb{R}^2$ be a domain on which a grayscale image $I : X \rightarrow [0, 1]$ is defined. The image can be fully represented as a collection of its *level sets* $\{x \in X : I(x) = t\}$ for $t \in [0, 1]$. Topologically, a level set may contain zero or more connected components of dimension 0 (points) or 1 (isolines).

Thinking of t as time and observing the evolution of the level sets over time, we will see connected components appear, split, change genus, join, and disappear. The study of the changes of topology of the level sets with infinitesimal changes of t belongs to the domain of Morse theory, a branch of differential topology.

The image can be represented by its *contour* or *component graph*, in which 1) a leaf vertex represents the creation or deletion of a component, 2) an interior vertex represents the joining/splitting of two or more components, and 3) an edge formed by two vertices with $t = t_1$ and $t = t_2$ represents a component in the level sets for all values of $t_1 \leq t \leq t_2$. This graph recording the topological events in the level set evolution can be shown to be a tree. Each edge of the component tree represents the evolution of a single connected component in some contiguous range of values of $t \in [t_1, t_2]$. We will denote such components by ∂R_t , implying the entire sequence $\{R_t\}_{t=t_1}^{t_2}$; $\text{int}(R_t)$ will denote the open set in X enclosed by ∂R_t , and R_t will denote the union of the two (the region with its boundary). Components R_t along the edge are *nested* inside each other.

2.1 Curvature Flow and Geometric Scale Space

In the SIFT method, interesting feature points are located by looking for local maxima of the discrete image Laplacian at different scales obtained by convolving the image with Gaussians of different variances. This procedure is known as *linear scale-space* analysis, and is equivalent to applying a linear diffusion equation $I_t = \Delta I$ on the image and observing the result at different times. While being scale invariant, the linear scale space is not affine invariant, and is thus unsuitable for images of the same scene captured from different viewpoints. Moreover, it is well known that linear scale space does not necessarily simplify the image structure. This is especially acute when level sets are considered, as linear scale space can disconnect simply connected shapes [11], [12].

Better scale-invariant quantities that are simplified with scale are provided by the *curvature scale space* or its affine variations [13], [14], [15], [16], [17], [18]. The construction of such a geometric scale space involves a nonlinear diffusion equation, which is more demanding computationally. We would therefore like to use the structure provided by geometric scale space without explicitly computing it, a property that was trivially accomplished for the linear scale space.

In the construction of the curvature scale space of an image, the image level sets are propagated by their curvature vector. Let



Fig. 1. First row: Random affine transformations of the silhouettes of the Puma logo (left) and a boy (right). Second row: Normalized shapes with second-order moments. Third row: Results of alternative method proposed by Cao et al.

$C(s) : [0, L] \rightarrow \mathbb{R}^2$ be an arclength-parameterized contour. Then, the curvature flow for the contour is given by

$$C_t(s) = C_{ss}, \quad (1)$$

where $C_{ss} = \kappa \vec{n}$ is the curvature vector, normal to the curve at $C(s)$. The whole process can be evaluated simultaneously for all the level sets using a nonlinear diffusion equation:

$$I_t = \operatorname{div} \left(\frac{\nabla I}{\|\nabla I\|} \right) \|\nabla I\|, \quad (2)$$

which can be easily established by the Osher-Sethian level set formulation [16]. The remarkable property of this flow proven by Grayson [14] is that embedding is preserved along the curvature flow and no self-intersections occur until the contour vanishes at a circular point.² Another important property is that each level set contour vanishes at a time proportional to its area at $t = 0$ [15], [14]. Thus, the curvature scale space can be used to define a component tree, capturing the topology of this scale space.

3 AFFINE-INVARIANT FEATURE DETECTION

A typical feature-based image analysis application (e.g., stereo matching) involves the steps of detecting regions in the image and describing them. The first stage of feature detection should be ideally repeatable in different views of the same scene, which is provided if the detection process is *affine covariant*, i.e., commutes with the geometric transformation of the image (in Section 4.1, we argue that this is not enough if blur is present in the scene). A feature descriptor is then extracted from the detected regions, which are normalized in order to undo the affine transformations. In this section, we briefly overview these two stages.

3.1 Maximally Stable Extremal Regions

Let R_t be the family of connected components representing an edge in the component tree. Matas et al. [4] refer to such regions as to *extremal* since either $I|_{\operatorname{int}(R_t)} < I|_{\partial R_t}$ or $I|_{\operatorname{int}(R_t)} > I|_{\partial R_t}$, i.e., all of the pixel values in the regions are either strictly darker or strictly brighter than those on the boundary, where the intensity is exactly equal to t .

The *stability* of a region R_t is defined as

$$\Psi_1(R_t) = \frac{A(R_t)}{\frac{d}{dt} A(R_t)}, \quad (3)$$

where $A(R_t)$ denotes the area of R_t . A region is considered stable if its area changes only slightly with the change of the threshold t . A region R_t is called *maximally stable* if $\Psi_1(R_t)$ has a local maximum at t (in practice, often computed over an interval of t). Such regions are image features detected by the MSER algorithm.

2. For a detailed discussion of affine invariants and scale-space construction, the reader is referred to chapters 2.2 and 3.1 in [19].

3.2 Affine-Invariant Normalization

In typical applications, maximally stable regions found by MSER undergo a process of affine-invariant normalization or canonization [20]. Normalization can be thought of as a mapping $N : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ receiving a region R and returning another region $N(R)$ such that $N(TR) = N(R)$ for any affine transformation T . Canonization of a given shape can be viewed as part of a descriptor computation in which the goal is to compensate for arbitrary transformations of the shape due to the acquisition process.

One of the easiest methods is *moment-based* canonization, in which the matrix of second-order geometric moments of the shape is diagonalized. Cao et al. [6] argue that such a normalization can be unstable, and propose alternatives based on the detection of flat intervals along the boundary. The next steps applied by Cao et al. involve center of mass estimation for the two regions created by a line parallel to the flat boundary line that goes through the center of mass. Parallel lines, area ratio, and center of mass are indeed robust measures preserved by affine transformations. On the other hand, a definition of flatness that is based on euclidean distance and angles is not invariant to affine transformations. Moreover, if we limit our discussion to the analysis of simple closed contours, there is a simple alternative for the first step proposed in [6].

Experimenting with second-order moments-based normalization [21], we did not experience the instabilities reported by Cao et al.: Moments-based normalization proved to be equally stable as the centers of mass-based alternative, as can be seen in Fig. 1.

A different normalization method proposed in [22] could be used to either initialize the Cao et al. canonization method or as compensation for the rotation ambiguity in moments-based normalization. Let us assume that the contours we would like to normalize are in general nonconvex (convex contours can be approximated by simple regular polygons). Relying on area ratios and centers of mass, and based on [6], we define a robust method for shape normalization (Algorithm 1; see also Figs. 2 and 3) allowing to undo affine transformations.

Algorithm 1. Shape normalization.

Input: Binary shape

Output: Normalized binary shape

- 1 Compute the convex hull of the shape.
- 2 Find the largest area bounded between the convex hull and the shape, and detect the bitangent line (part of the convex hull touching the largest area, as shown in Figure 2).
- 3 Follow the rest of the steps in [6] using the computed bitangent as the reference axis (Figure 3).

The reference axis could also be used for compensating for rotation ambiguity in the case of moments-based normalization



Fig. 2. Left to right: The shape's boundary contour, its convex hull, and the areas formed between the convex hull and the shape. The largest area, A_1 in this case, defines the bitangent that is used for normalization (canonization) of the shape or for fixing its orientation.

[21]: First, the normalization is performed, and then the above rotation cancellation using the convex hull and maximal bounded area is applied.

There are other options to account for rotations, like radial Fourier transform over the shape and consideration of the phase as a rotation angle. Yet, the best computational complexity for the convex hull of a closed contour is $O(n \log h)$, where h is the number of points in the convex hull ($n > h$) [23], while the Fourier transform is slightly costlier and requires $O(n \log n)$ operations.

4 INTERESTING FEATURES

The main focus of this section is a critique of the region stability criterion Ψ_1 used in the MSER feature detector [4]. The two main drawbacks we stress are the preference for rounded regions, which are of less interest in typical images, and the lack of affine invariance in the presence of blur. Understanding these drawbacks, we propose different stability criteria that are more robust and detect more interesting regions.

4.1 Stability and Shape Factor

Observing that

$$\frac{dA(R_t)}{dt} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (A(R_{t+\epsilon}) - A(R_t)) = \int_{\partial R_t} \frac{ds}{\|\nabla I\|}, \quad (4)$$

we can rewrite the MSER stability criterion Ψ_1 as

$$\Psi_1(R_t) = \frac{A(R_t)}{\int_{\partial R_t} \frac{ds}{\|\nabla I\|}}. \quad (5)$$

Let us now apply Ψ_1 to two equal-area regions, one is a perfect circle, while the other is a more interesting less round shape. Under the simplifying assumption that the change of intensity along the boundaries is the same in both regions, say $\|\nabla I\| = 1$, we have

$$\Psi_1(R_t) = \frac{A(R_t)}{\int_{\partial R_t} ds} = \frac{A(R_t)}{L(\partial R_t)}, \quad (6)$$

where $L(\partial R_t)$ is the boundary length of R_t . Similar to the *shape factor* $\frac{4\pi A}{L^2}$ which is always smaller or equal to 1 with equality achieved for the circle, the ratio $\frac{A}{L}$ grows when L decreases. Thus, the stability criterion Ψ_1 prefers regular shapes: For two equal-area shapes with the same change of intensity along their boundaries, the one with a shorter boundary will result in a larger Ψ_1 . However, such shapes are not necessarily the most interesting and descriptive features in natural images; on the contrary, interesting features typically have irregular boundaries.

Based on this observation, our goal is to correct the bias of Ψ_1 toward round shapes and define an alternative stability measure that prefers less regular and more interesting shapes while still enjoying the affine invariance and stability of Ψ_1 . Unlike the standard MSER, where Ψ_1 is computed on the components from the component tree, we propose to first normalize each component. We then compute the inverse of the standard euclidean shape factor

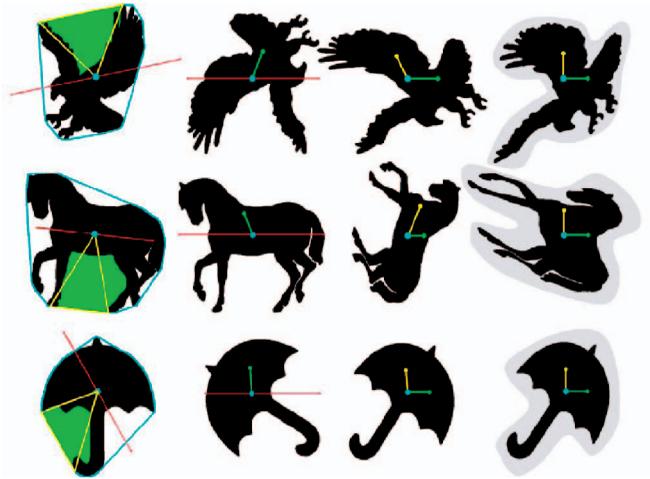


Fig. 3. Normalization steps of a given shape, left to right: Convex hull and maximal bounded area detection, rotation of the parallel to the bitangent through the center of mass, alignment of the center of mass of the upper half of the shape with the x -axis, and, finally, shear of the center of mass of the (new) upper part so that the line connecting it to the center of mass aligns with the y -axis. The resulting normalized shape is at the right of each sequence.

$$\Psi_2(R_t) = \frac{L^2(N(R_t))}{A(N(R_t))}, \quad (7)$$

where the operator N means that the measure is applied to the normalized region. Such a function prefers shapes with irregular boundaries, while still being affine invariant. Since the computational complexity of region normalization is proportional to the length of the boundary, reversing MSER selection and normalization is not more computationally expensive than first computing Ψ_1 and then doing normalization of the remaining MSERs.

It is possible to combine the above measure with Ψ_1 , e.g., $\Psi_1 \cdot \Psi_2$. Such a combination shows better performance, as demonstrated in the Section 5. Alternatively, it is possible to detect MSERs as extrema of Ψ_1 , but score them using Ψ_2 .

4.2 Affine Covariance and Blur

Matas et al. [4] showed that MSER is *affine covariant*. This observation stems directly from the fact that area ratios are preserved under affine transformations, which implies that $\Psi_1(R_t)$ is an affine-invariant property. This, in turn, implies that for an affine transformation T of the domain X , the corresponding regions R and R' detected in images I and $I(T^{-1})$, respectively, are related by $TR = R'$.

Affine covariance of maximally stable regions is the consequence of covariance of the level sets of the image with affine transformations of the coordinates. However, this property holds only if the boundaries of objects in the scene are smooth, an assumption far from true in real-world scenarios. Specifically, in order for the affine covariance of the level sets to hold, we need the optical point spread function of the camera to be small compared to the natural smoothness of objects in the scene. In other words, this leads to the assumption that the world is *first* blurred, and that the image formation is primarily a geometric transformation of that blurred image of the world. In a more realistic model, the blur occurs *after* the geometric transformation, i.e., real viewpoint transformations constitute (locally) affine transformations followed by blur in the image plane with the point spread function of the camera, and these two processes do not commute (Fig. 4).

As in most practical cases, the image formation involves nonnegligible blur due to the optical acquisition process; it may happen that the criterion Ψ_1 is not truly invariant to viewpoint transformations. A somewhat better quantity for the stability or edginess of a region would be the weighted gradient magnitude

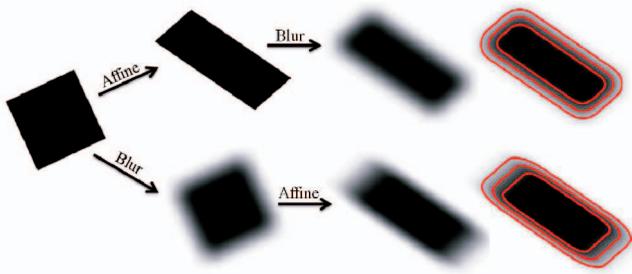


Fig. 4. The top row assumes affine transformation followed by imaging blur. The bottom row assumes affine transformation of a given blurred object. On the right are three corresponding level sets for both cases.

TABLE 1
Repeatability (in Percent) of MSER Feature Detector with Different Stability Criteria on the *Graffiti Wall* Data Set

Transformation	Stability criterion		
	Ψ_1	Ψ_2	$\Psi_1 \cdot \Psi_2$
View 20^0	60.3	58.5	58.0
View 30^0	57.4	58.3	58.1
View 40^0	40.8	49.2	47.0
View 50^0	32.6	36.9	32.5
View 60^0	21.1	27.5	25.8

along its boundary (similar criteria are used in edge detection [24], [25]). Here, in order to preserve affine invariance, the affine arclength $dv = |\kappa|^{1/3} ds$ can be used, which explicitly yields

$$\Psi_3(R_t) = \frac{A(R_t)}{\int_{\partial R_t} \frac{\|I_{xx}I_y^2 - 2I_xI_yI_{xy} + I_{yy}I_x^2\|^{1/3} ds}{\|\nabla I\|}}. \quad (8)$$

Any alternative robust filter, like the median, could represent the significance of the boundary sufficiently well.³ Such a Ψ_3 can be used instead of Ψ_1 .

5 EXPERIMENTAL RESULTS

To experimentally assess the findings of this paper, we closely followed the evaluation of Mikolajczyk et al. [9] for affine-invariant region detectors. As a reference MSER implementation, we used the publicly distributed VLFeat code of Vedaldi and Fulkerson [28], in which the stability criterion was modified according to the paper. The settings used were $\delta = 6$, minimum diversity of 0.4, maximum and minimum region area of 75 percent of the image size and 40 pixels, respectively (for additional details on the parameters, see [28]). Moment-based affine normalization as in Section 3.2 was used. We provide a quantitative comparison for stability criteria Ψ_1 , Ψ_2 , and $\Psi_1 \cdot \Psi_2$, and a qualitative evaluation of Ψ_3 .

5.1 Repeatability

In the first experiment, we tested the repeatability of the feature detector on a set of images with known geometric transformations using the methodology and code from [9]. Given a pair of such images, MSER regions represented as ellipses are detected in each of them independently. The repeatability score for a given pair of images is computed as the ratio between the number of region-to-region correspondences known from the geometric transformation between the images and the smaller of the number of regions in the

3. Note that the two basic independent affine-invariant second-order differential descriptors are $J(I) = I_{xx}I_y^2 - 2I_xI_yI_{xy} + I_{yy}I_x^2$, and the determinant of the hessian $H(I) = I_{xx}I_{yy} - I_{xy}^2$ [26], while the second-order approximation for the affine-invariant curvature of the level sets is given by $\mu = H/J^{2/3}$ [27].

TABLE 2
Repeatability (in Percent) of MSER Feature Detector with Different Stability Criteria on the *Blurred Graffiti Wall* Data Set

Transformation	Stability criterion		
	Ψ_1	Ψ_2	$\Psi_1 \cdot \Psi_2$
View 30^0 , no blur	57.4	58.3	58.1
View 30^0 , blur 1.5	32.6	41.1	42.7
View 30^0 , blur 2	27.6	36.0	36.6
View 30^0 , blur 2.5	21.4	33.2	30.7
View 30^0 , blur 3	21.4	32.9	27.5

TABLE 3
Matching Score (in Percent) of MSER Feature Detector with Different Stability Criteria on the *Blurred Graffiti Wall* Data Set

Transformation	Stability criterion		
	Ψ_1	Ψ_2	$\Psi_1 \cdot \Psi_2$
View 30^0 , no blur	34.7	25.6	29.2
View 30^0 , blur 1.5	10.0	11.7	13.2
View 30^0 , blur 2	6.5	6.5	10.9
View 30^0 , blur 2.5	4.8	5.8	7.5
View 30^0 , blur 3	4.1	4.7	5.9

pair of images (only the regions located in the part of the scene present in both images are considered). Regions are declared corresponding if their overlap error is below 50 percent [9]. Ideal repeatability is 100 percent.

The first data set (*graffiti wall*) included six images from the Mikolajczyk et al. benchmark [9], taken at different view angles from mild (20 degrees) to extreme (60 degrees). Table 1 summarizes the repeatability of the MSER feature detector with different stability criteria discussed in the paper. The second data set (*blurred graffiti wall*) emulated the scenario depicted in Fig. 4 (top). It included two viewpoints of the graffiti wall: 0 degree used as reference and 30 degrees, followed by Gaussian blur of different variance ($\sigma = 0, 1.5, 2, 2.5, 3$ pixels). The repeatability results for this data set are shown in Table 2. In both settings, the proposed criteria (Ψ_2 and $\Psi_1 \cdot \Psi_2$) outperform the original MSER criterion (Ψ_1) in the presence of strong transformations. Moreover, the performance drop of Ψ_1 (2.85 times from mildest to strongest transformation in the *graffiti wall* data set and 2.68 times on *blurred graffiti wall*, respectively) is significantly larger than that of Ψ_2 (2.12 times on *graffiti wall* and 1.77 times on *blurred graffiti wall*) and $\Psi_1 \cdot \Psi_2$ (2.24 times on *graffiti wall* and 2.11 times on *blurred graffiti wall*).

5.2 Feature Matching

In the second experiment, we evaluated the feature matching quality using the methodology and code from [9]. Each detected MSER region was normalized to undo the affine transformation and represented as a 41×41 image. Then, the SIFT descriptor was computed in the normalized regions. Closest features in the descriptor space were matched and the matching score was computed as the ratio of correctly matched regions and the number of corresponding regions obtained with the ground truth. Matching scores computed on the blurred graffiti wall data set are shown in Table 3. The matching of MSERs detected using the proposed stability criteria shows better robustness to view and blur transformations. As in the first experiment, the performance drop of Ψ_2 (5.4 times from mildest to strongest transformation) and $\Psi_1 \cdot \Psi_2$ (4.94 times) is smaller than that of Ψ_1 (8.4 times).

Finally, in the third experiment, we provide a qualitative comparison of Ψ_1 and Ψ_3 . Fig. 5 shows a different example of feature matching in an object taken from two video frames of a movie. The MSER regions are normalized and matched based on



Fig. 5. Feature matching in two frames of a video using classical MSER Ψ_1 (left) and our stability criterion Ψ_3 (right).

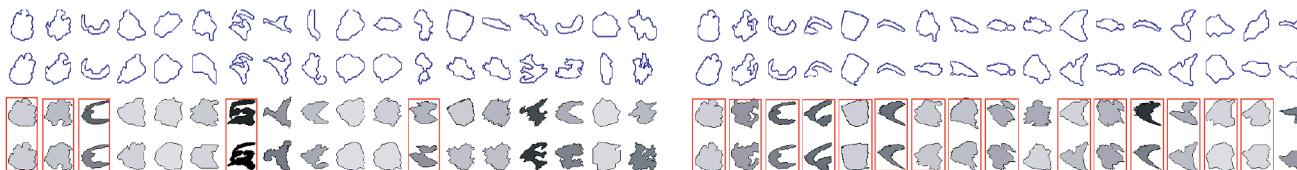


Fig. 6. Matching feature pairs extracted with the classical MSER Ψ_1 (left) and our stability criterion Ψ_3 (right). First and second rows: Regions found in the first and second frames. Third and fourth rows: Normalized regions in the first and second frames. The order (left to right) is according to the matching score, while the gray level of the canonical shapes corresponds to the isoperimetric ratio. Correct matches are boxed.

their canonized shapes, and for each pair, the first three matches are considered. The final selection is of features that are supported by consistent neighboring features that are determined by the first 10 nearest neighbors. The improvement in matching performances shows up in the correspondence of features in the two frames as can be seen in Fig. 6.

6 CONCLUSIONS

In this paper, we analyzed the MSER feature detection algorithm. We revisited the assumptions of the MSER and redefined some of the criteria that help us extract more informative shape descriptors using the curvature scale-space formalism. We stress again the amazing fact that while being only euclidean invariant, the curvature scale-space structure is captured by the level set graph, which is affine invariant. This property explains the usefulness of the image level sets and their local density in generating interesting features. The relation between the level set graph, curvature flow, and invariant stable and interesting features provides a theoretical bridge that could be used for various image and shape analysis applications.

ACKNOWLEDGMENTS

This research was supported by the European Commission's FP7-ERC program, grant agreement no. 267414. Michael M. Bronstein is supported by the Swiss High-Performance and High-Productivity Computing (HP2C) grant.

REFERENCES

- [1] J. Sivic and A. Zisserman, "Video Google: A Text Retrieval Approach to Object Matching in Videos," *Proc. IEEE Int'l Conf. Computer Vision*, 2003.
- [2] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int'l J. Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [3] G. Yu and J.M. Morel, "A Fully Affine Invariant Image Comparison Method," *Proc. IEEE Int'l Conf. Acoustics, Speech and Signal Processing*, pp. 1597-1600, 2009.
- [4] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust Wide Baseline Stereo from Maximally Stable Extremal Regions," *Proc. British Machine Vision Conf.*, pp. 384-393, 2002.
- [5] A. Desolneux, L. Moisan, and J.M. Morel, "Edge Detection by Helmholtz Principle," *J. Math. Imaging and Vision*, vol. 14, pp. 271-284, 2001.
- [6] F. Cao, J. Lisani, J.M. Morel, P. Musé, and F. Sur, *A Theory of Shape Identification*. Springer, 2008.
- [7] A.M. Bruckstein, R.J. Holt, A. Netravali, and T.J. Richardson, "Invariant Signatures for Planar Shape Recognition under Partial Occlusion," *CVGIP: Image Understanding*, vol. 58, no. 1, pp. 49-65, 1993.
- [8] A.M. Bruckstein and D. Shaked, "Skew Symmetry Detection via Invariant Signatures," *Pattern Recognition*, vol. 31, no. 2, pp. 181-192, 1998.

- [9] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. van Gool, "A Comparison of Affine Region Detectors," *Int'l J. Computer Vision*, vol. 65, nos. 1/2, pp. 43-72, 2005.
- [10] P. Forssén and D. Lowe, "Shape Descriptors for Maximally Stable Extremal Regions," *Proc. IEEE Int'l Conf. Computer Vision*, 2007.
- [11] *Geometric-Driven Diffusion in Computer Vision*, B.M. ter Haar Romeny, ed. Kluwer Academic Publishers, 1994.
- [12] F. Guichard and J.M. Morel, "Image Analysis and P.D.E.s," *IPAM GBM Tutorial*, 2001.
- [13] B.B. Kimia, "Toward a Computational Theory of Shape," PhD dissertation, Dept. of Electrical Eng., McGill Univ., 1990.
- [14] M.A. Grayson, "The Heat Equation Shrinks Embedded Plane Curves to Round Points," *J. Differential Geometry*, vol. 26, pp. 285-314, 1987.
- [15] M. Gage and R.S. Hamilton, "The Heat Equation Shrinking Convex Plane Curves," *J. Differential Geometry*, vol. 23, pp. 69-96, 1986.
- [16] S.J. Osher and J.A. Sethian, "Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations," *J. Computational Physics*, vol. 79, pp. 12-49, 1988.
- [17] G. Sapiro, "Topics in Shape Evolution," DSc thesis, Technion-Israel Inst. of Technology, 1993.
- [18] L. Alvarez, F. Guichard, P.L. Lions, and J.M. Morel, "Axioms and Fundamental Equations of Image Processing," *Archive for Rational Mechanics and Analysis*, vol. 123, pp. 199-257, 1993.
- [19] R. Kimmel, *Numerical Geometry of Images*. Springer, 2004.
- [20] Š. Obdržálek and J. Matas, "Object Recognition Using Local Affine Frames on Maximally Stable Extremal Regions," *Toward Category-Level Object Recognition*, pp. 85-108, Springer, 2006.
- [21] M. Hu, "Visual Pattern Recognition by Moment Invariants," *IEEE Trans. Information Theory*, vol. 8, no. 2, pp. 179-187, Feb. 1962.
- [22] R. Kimmel, C. Zhang, A.M. Bronstein, and M.M. Bronstein, "Feature-Based Image Alignment via Coupled Hough Transforms," Technical Report 2009-09, Technion, CIS, 2009.
- [23] D.G. Kirkpatrick and R. Seidel, "The Ultimate Planar Convex Hull Algorithm," *SIAM J. Computing*, vol. 15, no. 1, pp. 287-299, 1986.
- [24] P. Fua and Y. Leclerc, "Model Driven Edge Detection," *Machine Vision and Applications*, vol. 3, no. 1, pp. 45-56, 1990.
- [25] R. Kimmel and A.M. Bruckstein, "On Edge Detection, Edge Integration and Geometric Active Contours," *Proc. Int'l Symp. Math. Morphology*, p. 37, 2002.
- [26] P.J. Olver, *Equivalence, Invariants, and Symmetry*. Cambridge Univ. Press, 1995.
- [27] R. Kimmel, "Affine Differential Signatures for Gray Level Images of Planar Shapes," *Proc. Int'l Conf. Pattern Recognition*, 1996.
- [28] A. Vedaldi and B. Fulkerson, "VLFeat: An Open and Portable Library of Computer Vision Algorithms," <http://www.vlfeat.org/>, 2008.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.