TRANSFORMATION-INVARIANT EXTRACTION OF MULTI-LOCATION IMAGE FEATURES FROM REMOTE SENSING IMAGERY

Roman M. Palenichka, Ahmed Lakhssassi and Marek B. Zaremba

Dept. of Computer Science and Engineering
Université du Québec en Outaouais
Gatineau, Québec, Canada

1. INTRODUCTION

The feature extraction method proposed in this paper is developed in the context of object recognition, feature-based matching, image registration and context-based image retrieval. There are two major ways of image feature extraction: global approach with a single (and long) vector of global descriptors [1] and multi-location approach with a set (or a relationship graph) of local descriptor vectors [1, 2]. Since the image may contain many different objects of interest with different spatial extents (local scales), many local image features have to be extracted. The second approach is preferred in the practice of remote sensing because of its high descriptive power, relational (i.e., context) description, stability to distortions, and resistance to occlusions. An acceptable solution in the multi-location approach to feature extraction must satisfy three basic conditions: *relevance*, *invariance* to assumed image transformations and *robustness*.

A novel method of transformation-invariant feature extraction called *multi-location saliency pattern* (MSP) is proposed in this paper for object recognition and image matching. Multi-location image features are extracted in salient image points corresponding to local maxima of an attention operator, which indicates image locations with high intensity contrast, region homogeneity and shape saliency. A new and advantageous characteristic of the MSP is the incorporation of invariant planar shape descriptors into each descriptor vector, which also contains local pose parameters and intensity (color, multi-spectral) descriptors. Presence of a particular shape feature means local uniqueness that is a prerequisite in achieving feature relevance by the MSP method.

2. MULTI-SCALE DETECTION OF SALIENT POINTS

Existing attention operators including salient point detectors can be used to detect salient image locations for the MSP extraction. Salient point detectors, which are mostly represented by corner detectors, became popular in image feature extraction [3, 4]. However, the majority of the corner detectors are not scale-invariant that will result in detecting false salient points such a simple (straight-line) edges and other irrelevant image locations. On the other hand, existing multi-scale attention operators have poor performance at large scales, i.e., when salient points correspond to center points of relatively large homogeneous regions with high contrast. Therefore, we have used a modified version of the MIMF operator for the MSP extraction since it generally works well at large scales (region sizes) and can be implemented in a fast recursive way [5]. The salient points are extracted as consecutive local maxima of the MIMF operator in an image block A as follows:

$$(x,y)_p = \arg\max_{(i,j)\in A}\left\{\Phi[f(i,j),\rho(i,j)],(i,j)\notin \Gamma_{p-1}\right\}, \quad (1)$$

where $f(i,j)$ is the image intensity function, $\Phi[f(i,j),\rho(i,j)]$ is the MIMF operator function at the local scale $\rho(i,j)$, and $(x,y)_p$ are the coordinates of the pth local maximum. The region Γ_{p-1} in Eq. 1 is the masking region that excludes areas around the previously determined maxima from further extraction. The local scale value is a core element of the MIMF operator and estimated in each point of the image block A. According to the adopted morphological definition, the scale determination consists in selecting the greatest by diameter disk centered at (i,j) and inscribed into the current homogeneous region [5]. The scale value ρ is the diameter of the circular fragment of the homogeneous region.

3. CONCISE IMAGE DESCRIPTION BY MULTI-LOCATION SALIENCY PATTERNS

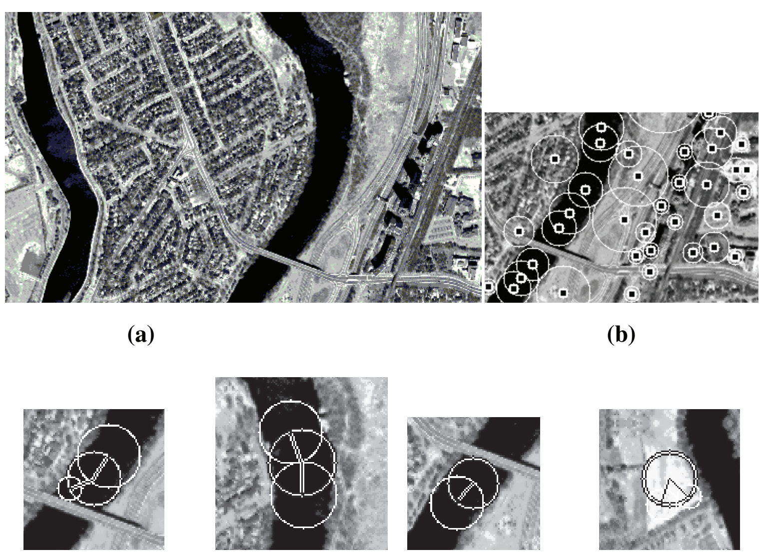
The underlying idea for the MSP method of image description is to represent the entire image or its block by a set of K local descriptor vectors $\{\xi_k\}$, where $\xi_{k,l}$ denotes lth descriptor in the kth descriptor vector related to kth salient point $(i,j)_k$. Each vector ξ_k consists of L heterogeneous scalar descriptors, which concisely describe one salient image fragment that is a circular neighborhood of $(i,j)_k$ with the diameter $\rho_k(i,j)$. Three distinctive types of fragment descriptors are extracted to form the descriptor vector ξ_k: 1) pose characteristics (i.e., two coordinates of salient point, local scale and local direction); 2) region shape descriptors [6]; 3) intensity (color or multi-spectral) parameters [3, 5]. They all are independent and normalized to be contained in the range of [0; 1]. The intensity parameters are transformation-invariant since they are estimated using an isotropic window averaging [2, 3]. The pose descriptors as well as region shape descriptors are made invariant to similarity transformations (i.e., translation, scaling and rotation). In particular, the transformation invariance of the pose descriptors is achieved by the descriptors estimation relatively to the *centroid, dominant scale and rotation angle* of a given MSP. The MSP centroid (X,Y) is

$$X = \frac{\sum_k \rho_k \cdot x_k}{\sum_k \rho_k} \quad \text{and} \quad Y = \frac{\sum_k \rho_k \cdot y_k}{\sum_k \rho_k}, \quad (2)$$

where x_k and y_k are the absolute coordinates on the image plane of kth salient point belonging to a given MSP. The dominant scale ρ_d of the MSP is the maximum local scale over all K salient points. The MSP rotation angle is determined by the angle of the line connecting the MSP centroid and the salient point that has the dominant scale ρ_d. Descriptor-invariance transformation of the pose descriptors then proceeds as follows. The salient point coordinates of each vector ξ_k are determined relatively (as the difference) to the MSP centroid in Eq. 2. The ratio of local scale to dominant scale ρ_d is assigned to the local scale value. Local direction descriptor in the vector ξ_k is the value of direction angle minus the MSP rotation angle.

4. REGION SHAPE DESCRIPTION USING LOCAL SKELETAL FEATURES

Planar shape descriptors play a key role in matching and recognition of objects of interest since they usually determine the uniqueness of image fragments containing object of interest. Intensity feature are subject to variability and may be completely different or absent at all while matching heterogeneous images. Skeletal shape is an economical way to describe homogeneous image regions in remote sensing, however invariance to similarity transformations and computational complexity remain an issue. We have modified local shape descriptors based on the piecewise-linear representation of skeletal shapes in order to incorporate them into the MSP concept [6]. The local region shape is a sequence of descriptor pairs [6]: angle of straight-line radial segments originating from the salient point and local scales at the segment vertices (Fig. 1 c-f). However, the direct use of the shape descriptors give poor results if comparing two local shapes with different numbers of straight-line segments. Therefore, we have adapted this method to the case of uneven number of segments to be used in the MSP approach. First, the segment angles are calculated relatively to the rotation angle of *k*th salient point (descriptor vector). Second, the descriptor pairs are sorted in the order of their local scale value and normalized with respect to the maximum local scale.



(a) (b)

(c) (d) (e) (f)

Fig. 1. Example of the MSP extraction (b) from image (a) and four examples of salient image fragments (c-f). Local scales at salient points are shown by white circles (the circle diameter is the local scale value).

5. EXPERIMENTAL RESULTS

The MSP method was tested on different examples of remote sensing imagery in order to experimentally prove the basic properties required for image feature extraction: relevance, invariance to image rotation and robustness. Figure 1 a shows an example of QuickBird image of an urban area and a MSP (Fig. 1b) extracted from an image block. Examples of salient image points (fragments) with their skeletal shape features are shown in Fig. 1 c-f.

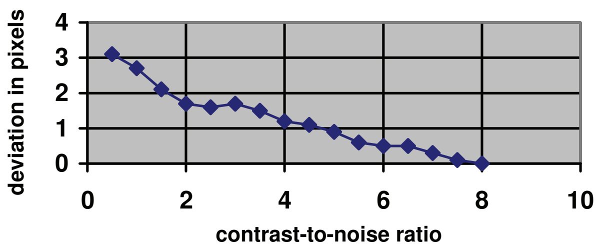


Fig. 2. Robustness of the local scale estimation in the MSP method.

The MSP descriptor invariance to rotation transformations was tested by rotating images on different rotation angles and comparing the obtained descriptor values with their reference values (i.e., at the zero rotation angle). The descriptors robustness against random distortions was tested by introducing synthetic noise to image fragments (model of white noise combined with image blurring) and measuring the deviations in the descriptor values versus different signal-to-noise ratio cases (example in Fig. 2). The overall test results have shown good performance of the MSP method for feature extraction and its high potential in object recognition and transformation-invariant matching of remote-sensing imagery.

Comparative experimental analysis of the MSP method of image description relatively to the existing SIFT (scale-invariant feature transform) method [7] has also been made. The proposed method has shown higher descriptive relevance and enhanced robustness of feature extraction due to the incorporated extraction of transformation-invariant skeletal shape features at larger scales (i.e., region sizes).

6. REFERENCES

1. T. Evgeniou, *et al.*, "Image representations and feature selection for multimedia database search", *IEEE Trans. KDE,* Vol. 15(4), pp. 911-920, 2003.
2. C. Schmid and R. Mohr, "Local gray-value invariants for image retrieval", *IEEE Trans. PAMI,* Vol. 19(5), pp. 530-535, 1997.
3. T. Kadir and M. Brady. "Saliency, scale and image description", *Int. J. of Comp. Vision*, 45(2)**,** pp. 85-103, 2001.
4. N. Sebe and M. S. Lew, "Comparing salient point detectors", *Pattern Recognition Letters*, Vol. 24(1-3), pp. 89-96, 2003.
5. R. M. Palenichka and M. B. Zaremba. "Multi-Scale Isotropic Matched Filtering for Individual Tree Detection in LiDAR Images"**,** *IEEE Trans. GRS***,** 45(12), pp. 3944-3957**,** 2007.
6. R. M. Palenichka and M. B. Zaremba, **"**Object shape extraction based on the piecewise linear skeletal representation", *Proc, Int. Conf. ICIAR2005,* Vol. LNCS-3656, pp. 464-472, 2005.
7. D. G. Lowe, "Distinctive image features from scale-invariant key-points", *Int. J. of Comp. Vision*, Vol. 60(2), pp. 91-110, 2004.