SCALAR IMAGE INTEREST POINT DETECTION AND DESCRIPTION BASED ON DISCRETE MORSE THEORY AND GEOMETRIC DESCRIPTORS

Ricardo Dutra da Silva[1], *William Robson Schwartz*[2], *Helio Pedrini*[1]

[1]Institute of Computing, University of Campinas, Campinas-SP, Brazil, 13083-852
[2]Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte-MG, 31270-010

ABSTRACT

The use of scalar data has arisen in many image applications which obtain data from simulated experiments, range scanners and photogrammetry. Their different nature of acquisition influences the type of processing and the analysis that may be undertaken in tasks such as registration, retrieval and recognition of structures or objects. This paper presents a new method for detecting and describing scale invariant descriptors over dense scalar data sets. The detection of interest points for each image is accomplished using an approach based on the Morse theory and descriptors are computed exploring the geometry of the data. Experiments are performed to demonstrate the effectiveness of the proposed method.

***Index Terms*—** Morse theory, interest points, geometric descriptor, scalar data

1. INTRODUCTION

Scalar data represent a great amount of image information used in areas such as remote sensing and simulation of phenomena. This kind of data is usually obtained by techniques different from those used for regular luminance images, which normally employ cameras. Scalar data are usually obtained by simulated experiments, range scanners and photogrammetry. Therefore, such data present a different nature that influences the type of processing and analysis that may be undertaken.

One important research topic related to images is the extraction of descriptors for tasks such as registration, retrieval and recognition of structures or objects. Descriptors as SIFT [1] have been successfully applied to these tasks on regular images. Nevertheless, gradient-based descriptors may not be as effective when applied to scalar data since their characteristics are different. Lodha and Xiao [2] consider this problem and propose a method combining SIFT-like interest point detection with a height-based local descriptor (opposed to SIFT gradient-based), which they refer to as the height difference histogram (HDH).

The detection of interest points is the subject of several methods such as Harris, Harris-Laplace, Hessian-Laplace and Difference of Gaussians (DoG) [3]. DoG is a common method for approximating the computation of the second derivative, which is extensively used in practice [4]. Nevertheless, identifying interest points with DoG requires much work for tuning parameters, such as mask size and thresholds, in search of stability.

Our assumption is that there are better choices for identifying important points over scalar data. The Morse theory relates the topology of a function with critical points (maxima, minima and saddles) [5, 6]. Critical Morse points are usually used to analyze n-dimensional scalar data taking into account the topological information [7].

In this paper, we propose a new method for detecting and describing scale invariant descriptors over dense scalar data sets. The detection of interest points through scales of a pyramid of Gaussians is based on critical points according to the Morse theory, which we conjecture to be more suitable for scalar data since it does not require tuning several parameters as in the DoG approach. Our method employs the discrete Morse theory because it works directly on the discrete data.

We compared our method against the DoG approach to demonstrate its effectiveness. We also propose an extension of Lodha's [2] height difference histogram, which we show to be more discriminative as a geometric descriptor.

The text is organized as follows. Section 2 presents related works and concepts. The proposed method is described in Section 3. Experimental results and comparisons are presented in Section 4. Finally, Section 5 concludes the paper with final remarks.

2. BACKGROUND

This section briefly reviews some important research and concepts related to the topic under investigation.

2.1. Pyramids of Gaussians

A pyramid of images is a sequence of images $\{I_1, I_2, \ldots, I_n\}$, where I_1 is the original image itself and I_{i+1} is derived from I_i by reducing its resolution by a factor $d \geq 1$, commonly 2 [8]. Let $\downarrow_d$ denote the interpolation operator to reduce resolution, then

$$I_{i+1} = \downarrow_d I_i. \tag{1}$$

The convolution of an image $I(x, y)$ with a family of Gaussian kernels $G(x, y, \sigma_i)$ using different standard deviation values σ_i produces a sequence of smoothed images $\{L_1, L_2, \ldots, L_n\}$, such that

$$L_i = L(x, y, \sigma_i) = G(x, y, \sigma_i) * I(x, y) \tag{2}$$

where

$$G(x, y, \sigma_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-(x^2 + y^2)/2\sigma_i^2}. \tag{3}$$

The pyramid of Gaussians, as defined in [9], is a multi-scale image representation built as a family of images $P_i(x, y)$, $i = 1, 2, \ldots, n$, computed by subsequent applications of a Gaussian kernel followed by an interpolated reduction of the resolution of the

Thanks to FAPESP, FAPEMIG, CNPq and CAPES for the financial support.

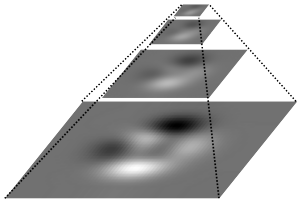

Fig. 1: Pyramid of Gaussians with four levels of resolution.

resulting image by a factor d. We can define the i-th image of the pyramid of Gaussians as

$$P_i = \begin{cases} G(x,y,\sigma) * I_1(x,y) & \text{if } i = 1 \\ G(x,y,\sigma) * (\downarrow_d P_{i-1}(x,y)) & \text{otherwise} \end{cases} \quad (4)$$

for a given fixed σ. An example of a pyramid of Gaussians is shown in Figure 1.

2.2. Difference of Gaussians

The second derivative of an image function $I(x,y)$ is obtained by the Laplace operator, which can be effectively approximated by convolving the image with a mask that is the difference of two Gaussian masks with substantially different values of σ [10]. Figure 2 shows an example of the Difference of Gaussians with two levels of resolution.

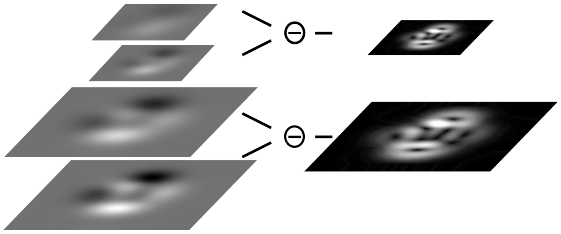

Fig. 2: Pyramid of Difference of Gaussians with two levels of resolution. Each level is formed by the subtraction of images smoothed by Gaussian masks with different values of σ.

The pyramid of DoG has been used for detection of interest points in images [1]. According to Lowe [1], given three images in the pyramid, a point is considered a local minima (maxima) if its value is smaller (greater) than any value inside a local cubical region properly located in the three images.

2.3. Discrete Morse Theory

The discrete Morse theory, formulated by Forman [5], is an adaptation of the Morse theory [11] to discrete structures. This theory relates the topology of a smooth function f with its critical points: maxima, minima and saddles.

Cell complexes are used as discrete structures for the discrete Morse theory. A p-cell σ^p of dimension p is a set of points which is homeomorphic to a closed unit ball B^p of dimension p. Cells of a complex are connected by boundary relations. The boundary of a p-cell, $\partial\sigma^p$, is the portion of the cell which is mapped by homeomorphism to the boundary of the unit ball. A cell τ is called a face of a p-cell, σ, if it is part of the boundary of the p-cell. As a consequence, σ is coface of τ. A cell complex K is a finite collection of cells such that all the faces of a cell in the complex also belong to the complex and the intersection of any two cell is either empty or a face of both cells.

A function $f\colon K \to \mathbb{R}$ is a discrete Morse function if for all $\sigma \in K$, f takes a value less or equal to $f(\sigma)$ in at most one coface of σ and takes a value greater or equal to $f(\sigma)$ in at most one face of σ. Given a discrete Morse function, a cell $\sigma^p \in K^p$ is critical of index p if

$$\sharp\{\tau^{p+1} > \sigma \mid f(\tau) \leq f(\sigma)\} = 0 \quad (5)$$

$$\sharp\{\lambda^{p-1} < \sigma \mid f(\lambda) \geq f(\sigma)\} = 0. \quad (6)$$

such that $\sharp$ denotes set cardinality. In a two dimensional cell complex, minima are 0-cells, saddles are 1-cells and maxima are 2-cells. Figure 3 shows an example of quadrangular cell complex and its critical cells.

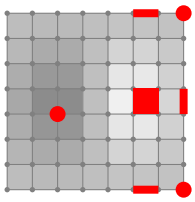

Fig. 3: Example of cell complex with critical Morse cells highlighted (minima are 0-cells, saddles are 1-cells and maxima are 2-cells).

Critical Morse cells can be identified as unmatched cells in a Morse matching [12]. Based on filtrations of a complex K [13] and simple homotopy theory [14], Robins et al. [13] recently developed an algorithm for identifying critical cells and building Morse complexes. We use their algorithm in our work.

3. PROPOSED METHOD

The proposed method for detecting and describing interest points is composed of four main steps: (1) computation of the Gaussian pyramid, (2) computation of the cell complex, (3) extraction of critical cells which will be used as interest points and (4) computation of descriptors for each interest point. These steps are summarized in Figure 4.

Initially, a pyramid of Gaussians $P_s(x,y)$, $s = 1,\ldots,n$, is computed using a decimation factor of 1.5 with bilinear interpolation. The second step comprises the computation of a cell complex over the scalar data. Although we have chosen to use triangular cells, quadrangular cells could be considered as well. Given the cell complex, the critical cells of the discrete Morse function, used as interest points, are detected with the algorithm described in Robins et al. [13]. The detection of these interest points takes into account the information of a growing space resulting from the filtration of the cell complex, the identification carries much more information about the data than the local detection accomplished by the DoG approach.

From the critical points detected, we use maxima and minima as interest points since they are stable. Even though saddle points present important topological information, they may introduce instability between corresponding points of scalar data. For this reason, they were not used in our method.

The descriptors associated with the interest points are based on the height difference histogram (HDH) presented by Lodha and Xiao [2]. Given an interest point p, the height difference of p and

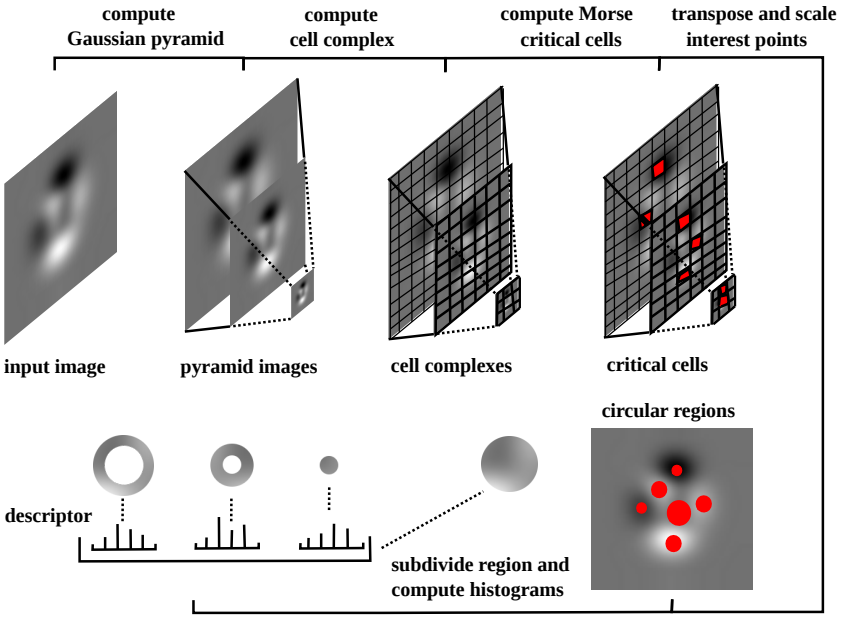

Fig. 4: Main steps of the proposed method. The input image is decomposed into a pyramid of Gaussians. A cell complex is computed using the pyramid and Morse critical cells are detected as interest points. Given a critical cell, a circular region is sampled with radius proportional to the scale in which the cell was detected. The descriptor is computed using radial subdivisions of the region.

a point q belonging to a region R around p is computed. The difference and its corresponding sign are used to index a bin of the histogram ranging from $-\Delta H$ to $+\Delta H$, where ΔH is the difference between the maximum and the minimum scalar values of the data set. The binning is computed at a log scale.

We use a different strategy to compute the height differences histogram: instead of using rectangular neighborhood regions, as in [2], the HDH of a specific interest point p, with coordinates (x, y), belonging to image P_s, is computed within a circular region around p. The circular region assures that the descriptors of two rotated patches of a same image locality are invariant. It is also reasonable to weight the data around an interest point so that a closer point will be more important than a farther one with similar height difference. A smooth function, such as a Gaussian or a distance-based function (used in our approach), would suffice.

Even though these considerations strength the descriptor, it is not uncommon to face similar histograms arising from regions that should have some dissimilarity. Then, we propose a subdivision of the circular region into m radial subregions to address this situation. Given a circular region of radius r, we subdivide this region into one circular region of radius r/m and $m-1$ rings formed by the regions between two consecutive circumferences of radius $k \times r/m$ and $(k+1) \times r/m$, $k = 1, \ldots, m-1$. Each subdivision creates a relative histogram. The concatenation of the histograms is the final descriptor of the interest point.

4. EXPERIMENTAL RESULTS

We applied our method to detect and describe terrain interest points in a registration task. The images used for the experiments were obtained from the set of elevation data available from the U.S. Geological Survey [15]. A set of ten images was used to create a data set of ninety pairs of reference and transformed images. Each pair corresponds to patches with 513×513 pixels randomly chosen from the original images. The transformed images were created applying rotation and reduction transformations with random parameters.

Our Morse-based interest point detection method was tested against the DoG approach. For both methods, we used the proposed extension of the HDH. The new descriptor was also evaluated to demonstrate that the subdivision yields good performance for terrain data when compared to the original descriptor.

The basic workflow for the image registration is composed of three steps: (1) matching of the descriptors extracted from interest points of the reference and transformed images; (2) extraction of pairs of line segments formed by pairs of interest points in the reference image and their counterparts matched in the transformed image; (3) estimation of the inverse transform parameters between all pairs of line segments.

The pyramids of Gaussians, as well as the pyramids of DoGs, were computed with a decimation factor of 1.5 and Gaussian mask with standard deviation $\sigma = 1.5$. Given the scale level s of the interest points in the pyramids of Morse or DoGs, experiments were carried out with circular regions of size $35 * (s-1) * 1.5$ if $s > 1$ and 35 if $s = 1$. The proposed descriptors were computed with $\Delta H = 45$ and three radial subdivisions of the circular region (see Section 3). The pyramids were computed with $s = 3$ and, therefore, transformed image rotation and reduction parameters were chosen inside the intervals $[0°, 360°]$ and $[1, 1.5^2] = [1, 2.25]$, respectively.

To achieve the matching of interest points, we used a one-against-all approach, computing the best match of each interest point of the transformed image among all interest points of the reference image. To perform the matching comparison between descriptors of interest points, we used the chi-squared distance, defined as

$$\chi^2(i,j) = \frac{1}{2} \sum_{k=1}^{n} \frac{(h_i(k) - h_j(k))^2}{h_i(k) + h_j(k)} \tag{7}$$

for histograms h_i and h_j of n bins each.

Given the matching, we define pairs of line segments between points p_i and p_j of the reference image and the matched points q_i and q_j of the transformed image. The line segments are used to estimate the parameters of rotation and scale of the inverse transform. After applying the estimation for all possible pair of lines, we take the mode of the estimated parameters to define the inverse transform.

In the first experiment, we computed the gain of subdividing the circular region used to create the descriptors. Some results are shown in Figure 5. All these tests were performed using our Morse-based method. The plot shows the percentage of interest points matched within a specific distance error compared to the ground-truth. We show in Figure 5 the results for regions of size $35 * (s-1) * 1.5$ and $45 * (s-1) * 1.5$ with and without subdivisions. The experiments were named M, x, y, where M stands for the Morse approach, x is the number of subdivisions of the circular region, and y is the radius of the circular region. The descriptors extracted from subdivided regions obtained a considerable gain in relation to the descriptors from non-subdivided regions.

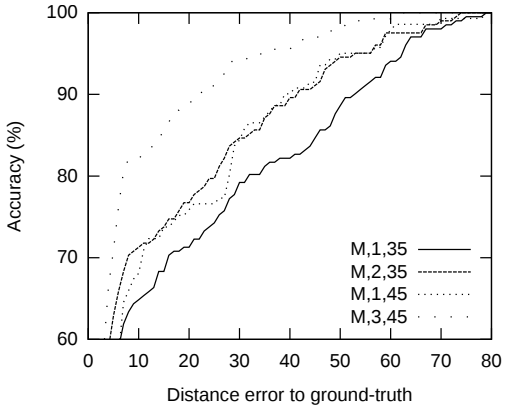

Fig. 5: Comparison of descriptor accuracy (percentage of correct matches as a function of the distance error) for non-subdivided and subdivided regions. The Morse-based approach $(M, 1, 35)$ with one region of size 35 without subdivision was less effective than the approach with region of size 35 and 2 subdivisions $(M, 2, 35)$. The same fact occurred with a 45 size region without subdivision $(M, 1, 45)$ and a 45 size region with three subdivisions $(M, 3, 45)$.

The second experiment considered the mean error of the estimated parameters for the inverse transform. The idea is to evaluate how accurate the detected points are. In a practical registration task, the RANSAC method could be considered to improve the outcomes. The results are summarized in Table 1. When using our descriptor both for Morse and Difference of Gaussian methods, the results were superior when compared to the approach with no subdivision (DoG2). The Morse-based method also achieved better performance when compared to the DoG method using the same settings for computing the descriptors. These results demonstrate that the Morse-based interest point detection is a promising alternative to DoG for scalar data. Although the Morse approach demands more computational time, it is worth mentioning that the detected number of points was in average less than that obtained by DoG approach, as shown in the fourth column of Table 1.

5. CONCLUSIONS

We presented a new method for detecting and describing scale invariant interest points over dense scalar data sets. The detection of interest points was accomplished using Morse critical maxima and minima, which was shown to be effective in comparison to the Difference of Gaussian method. Besides achieving promising results, the Morse-based approach avoids tuning several parameters and usually detects fewer interest points without losing accuracy. A second contribution of the work was the development of geometric descriptors, an extension of an earlier geometric descriptor, which improved the results for registration of terrain elevation data.

Approach	Rotation Error	Scaling Error	Number of Interest Points
Morse	0.10	1.11	639.94
DoG	0.16	1.40	1437.90
DoG2	0.16	1.84	1437.90

Table 1: Mean error of the resulting inverse transformations obtained with the tested methods.

6. REFERENCES

[1] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *IJCV*, vol. 60, no. 2, pp. 91–110, 2004.

[2] S. K. Lodha and Y. Xiao, "GSIFT: Geometric Scale Invariant Feature Transform for Terrain Data," in *Vision Geometry XIV*. SPIE, 2006, vol. 6066, pp. 1–11.

[3] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of Interest Point Detectors," *IJCV*, vol. 37, pp. 151–172, 2000.

[4] K. Mikolajczyk and C. Schmid, "A Performance Evaluation of Local Descriptors," *TPAMI*, vol. 27, pp. 1615–1630, 2005.

[5] R. Forman, "Morse Theory for Cell Complexes," *Advances in Mathematics*, vol. 134, no. 1, pp. 90–145, 1998.

[6] R. Forman, "Morse Theory and Evasiveness," *Combinatorica*, vol. 20, no. 4, pp. 489–504, 2000.

[7] A. Zomorodian, "Computational Topology," in *Algorithms and Theory of Computation Handbook*, M. Atallah and M. Blanton, Eds., vol. 2, chapter 3. Chapman & Hall/CRC Press, Boca Raton, FL, second edition, 2010.

[8] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*, Thomson Learning, 2007.

[9] T. Lindeberg, *Scale-Space Theory in Computer Vision*, Kluwer Academic Publishers, Netherlands, 1994.

[10] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, Henry Holt & Company, 1983.

[11] J. W. Milnor, *Morse Theory*, Annals of Mathematics Studies. Princeton University Press, 1963.

[12] M. Joswig and M. E. Pfetsch, "Computing Optimal Morse Matchings," *SIAM Journal on Discrete Mathematics*, vol. 20, no. 1, pp. 11–25, 2006.

[13] V. Robins, P. Wood, and A. P. Sheppard, "Theory and Algorithms for Constructing Discrete Morse Complexes from Grayscale Digital Images," *TPAMI*, vol. 33, no. 8, pp. 1646–1658, 2011.

[14] M. M. Cohen, *A Course in Simple-Homotopy Theory*, Springer, 1982.

[15] USGS, "National Elevation Dataset," 2011, `http://ned.usgs.gov/index.asp`.