Local 2D Pattern Spectra as Connected Region Descriptors

Petra Bosilj, Michael H.F. Wilkinson, Ewa Kijak, Sébastien Lefèvre

Local 2D pattern spectra as connected region descriptors

Petra Bosilj[1], Michael H. F. Wilkinson[2], Ewa Kijak[3], Sébastien Lefèvre[1]*

[1]Université de Bretagne Sud – IRISA, Vannes, France
[2]Johann Bernoulli Institute, University of Groningen, Groningen, The Netherlands
[3]Université de Rennes 1 – IRISA, Rennes, France

Abstract. We validate the usage of augmented 2D shape-size pattern spectra, calculated on arbitrary connected regions. The evaluation is performed on MSER regions and competitive performance with SIFT descriptors achieved in a simple retrieval system, by combining the local pattern spectra with normalized central moments. An additional advantage of the proposed descriptors is their size: being less than half the size of SIFT, they can handle larger databases in a time-efficient manner. We focus in this paper on presenting the challenges faced when transitioning from global pattern spectra to the local ones. An exhaustive study on the parameters and the properties of the newly constructed descriptor is the main contribution offered. We also consider possible improvements to the quality and computation efficiency of the proposed local descriptors.

Keywords: shape-size pattern spectra, granulometries, max-tree, region descriptors, CBIR

1 Introduction

Pattern spectra are histogram-like structures originating from mathematical morphology, commonly used for image analysis and classification [12], and contain the information on the distribution of sizes and shapes of image components. They can be efficiently computed using a technique known as granulometry [5] on a max-tree and min-tree hierarchy [9, 19].

We study here the application of 2D pattern spectra to Content Based Image Retrieval (CBIR), to retrieve database images describing the same object or scene as the query. Previous success in using the pattern spectra as image descriptors computed at the global [23, 24] or pixel scale (known as DMP [3] or DAP [6, 18]) convinced us to investigate their behavior as local descriptors.

Standard CBIR systems based on local descriptors consist of region detection, calculation of descriptors and storage in an index. Different indexing schemes are used to perform large scale database search [10, 22], but all need powerful local descriptors to achieve good performance [21]. To construct such a descriptor,

* The collaboration between the authors was supported by mobility grants from the Université européenne de Bretagne (UEB), French GdR ISIS from CNRS, and an excellence grant EOLE from the Franco-Dutch Network.

we want to extend [24] and compute 2D size-shape pattern spectra locally while keeping the good characteristics of the global version (scale, translation and rotation invariance, and computation efficiency). However, to evaluate the quality and properties of our proposed local pattern spectra (LPS) descriptors, we need to reexamine the parameters used with global pattern spectra as well as evaluate the effect of the new parameters introduced by the local descriptor scheme.

We evaluate our descriptors on the MSER regions [13] as they can also be computed on a max-tree [17], using the well-established SIFT descriptors [11] to obtain a baseline CBIR performance on a database. Future work will include comparisons with SIFT extensions which improve performance [1, 2]. A competitive precision is achieved with a rotation invariant version of the descriptor combined with normalized central moments, half the size of SIFT (deeper interpretation of the results and best achieved performance can be found in [4]).

As the goal of this paper is to give an overview of choices and challenges faced when reworking a global pattern spectrum into a local one, we adopt a slightly atypical presentation structure: The background notions are presented in Sec. 2, with the focus on how the max-tree is used in all parts of the CBIR system. The experimental framework used to tune and evaluate the descriptors is explained in Sec. 3. To examine the properties of the proposed LPS descriptor through the influence of parameters used, the main contribution can be found in Sec. 4, where the descriptor performance is also presented. Remarks on possible improvements to the efficiency of LPS computation are given in Sec. 5. Finally, the conclusions are drawn and directions for future work offered in Sec. 6.

2 Background

2.1 Max-tree

The concept of min and max-trees [9, 19] is here central for keypoint detection as well as the calculation of feature descriptors. We recall their definition using the *upper* and *lower level sets* of an image, e.g. sets of image pixels p with gray level values $f(p)$ respectively higher and lower than a threshold k.

Given a level k of an image I, each level set is defined as $\mathcal{L}^k = \{p \in I | f(p) \geq k\}$ for the max-tree, or $\mathcal{L}_k = \{p \in I | f(p) \leq k\}$ for the min-tree. Their connected components (also called the *peak components*) $\mathcal{L}^{k,i}$ and $\mathcal{L}_{k,i}$ (i from some index set) are nested and form a hierarchy. The min-tree is usually built as a max-tree of the inverted image $-I$.

2.2 MSER detection

Peak components of the upper and lower level sets $\{\mathcal{L}^{k,i}\}$ and $\{\mathcal{L}_{k,i}\}$ coincide with the maximal and minimal extremal regions in the context of *Maximally Stable Extremal Regions* (MSER) detector introduced by Matas et al. [13]. The detected regions correspond to bright and dark "blobs" in the image and can be extracted while building the max-tree and the min-tree [17].

Extraction of MSER relies on the stability function $q(\mathcal{L}^{k,i})$, which measures the rate of growth of the region w.r.t. the change of the threshold level k. It is

computed for all the elements of nested sequences, and the local minima of this function correspond to the maximally stable regions.

We use here a simplification commonly adopted by many computer vision libraries (e.g. VLFeat [26]) :

$$q(\mathcal{L}^{k,i}) = \frac{|\mathcal{L}^{k-\Delta,i} \setminus \mathcal{L}^{k,i}|}{|\mathcal{L}^{k,i}|}, \tag{1}$$

where the cardinality is denoted by $|\cdot|$ and Δ is a parameter of the detector. Additional parameters control the allowed region size, limit the appearance of too similar regions and impose a lower limit on the stability score.

2.3 Attributes and filtering

Region characteristics can be captured by assigning them *attributes* measuring the interesting aspects of the regions. *Increasing* attributes $K(\cdot)$ give increasing values when calculated on a nested sequence of regions, otherwise they are *nonincreasing*. A value of an increasing attribute on a tree region, $K(\mathcal{L}^{k,i})$, will be greater than the value of that attribute for any of the regions descendants.

Increasing attributes are usually a measure of the *size* of the region. We will simply use the *area* (in pixels) of the region, $A(\mathcal{L}^{k,i})$, as the size attribute. *Strict shape* attributes are the nonincreasing attributes dependent only on the region shape, thus invariant to scaling, rotation and translation [5]. To indicate the shape of a region, we use an elongation measure called *corrected noncompactness*:

$$NC(\mathcal{L}^{k,i}) = 2\pi \left(\frac{I(\mathcal{L}^{k,i})}{A(\mathcal{L}^{k,i})^2} + \frac{A(\mathcal{L}^{k,i})}{6} \right). \tag{2}$$

$I(\mathcal{L}^{k,i})$ is here the moment of inertia of the region, and the term $\frac{I(\mathcal{L}^{k,i})}{A(\mathcal{L}^{k,i})^2}$ without the correction is equal to the first moment invariant of Hu [8] $I = \mu_{2,0} + \mu_{0,2}$.

We will also directly use the normalized central moments $n_{1,1}, n_{2,0}, n_{0,2}, n_{0,4}$ and $n_{4,0}$ of the considered regions. These, and many more attributes (such as center of mass, covariances, skewness or kurtosis [27]) can be derived based on raw region moments.

When the tree is further processed by comparing the region attribute values to a threshold t (or using a more complex criterion), and making a decision to preserve or reject a region based on the attribute value, we are performing an *attribute filtering*. While filtering with an increasing attribute is relatively straightforward, advanced filtering strategies have to be used when performing a filtering with nonincreasing (e. g. shape) attributes [5, 19, 24].

2.4 Granulometries and global pattern spectra

Attribute opening is a specific kind of attribute filtering, in which the attribute used is increasing. Such a transformation is anti-extensive, increasing and idempotent. A *size granulometry* can be computed from a series of such openings, using increasing values for the threshold t. This series also satisfies the absorption property, since applying an opening with $t' < t$ will have no effect on an

Table 1: Subsets of the UCID database used in experiments.

	# categories / examples	categories selected
ucid5	31 / 5	all UCID categories with ≥ 5 examples
ucid4	44 / 4	all UCID categories ≥ 4
ucid3	77 / 3	all UCID categories ≥ 3
ucid2	137 / 2	all UCID categories ≥ 2
ucid1	262 / 1	all UCID categories

image already filtered with an opening using the threshold t. In other words, a size a granulometry can be seen as a set of sieves of increasing grades, each letting only details of certain sizes [24] pass through.

Instead of focusing on the details remaining, it is also possible to consider the amount of detail removed between pairs of consecutive openings. Such an analysis has been introduced by Maragos [12] under the name *size pattern spectra*. It can be seen as a 1D histogram containing, for each size class or filtering residue, its Lebesgue measure (i. e. the number of pixels in the binary case or the sum of gray levels in the grayscale case). Such histograms can also be computed over different shape classes, leading to the concept of a shape-spectra [24]. Finally, both shape and size pattern spectra can be combined to build shape-size pattern spectra [24]. A shape-size pattern spectrum is a 2D histogram, where the amount of image detail for the different size-shape classes are stored in dedicated 2D bins.

Previous work [23, 24] as well as our own experiments suggest that the lower attribute values carry more information. Thus, a logarithmic binning is used for both attributes, producing higher resolution bins for low attribute values. Let v be the attribute value for one of the attributes, N_b the total desired *number of bins* and m the *upper* bound for that attribute (which can be the maximal attribute value in the hierarchy, or a smaller value if we decide to ignore attribute values above a certain threshold). If the minimal value for the attribute is 1 (as with both area and the corrected noncompactness), the base for the logarithmic binning b, and the final bin c, are determined as:

$$b = \sqrt[N_b]{m}, \tag{3}$$

$$c = \lfloor \log_b v \rfloor \tag{4}$$

Enumerating the bins starting from 1, the i-th bin has the range $[b^{i-1}, b^i]$.

Connected pattern spectra are effectively calculated in a single pass over a max-tree [5, 24]. For every region, we calculate both the size attribute $v_1 = A(\mathcal{L}^{k,i})$ and shape attribute $v_2 = NC(\mathcal{L}^{k,i})$, and add the area of the region weighted by its contrast with the parent region δ_h to the spectrum bin $S(c_1, c_2)$. Before using the spectrum as a descriptor, we equalize the sums in the bins as $\sqrt[5]{S(c_1, c_2)}$. More information and discussion about the algorithm used to compute the descriptors is given in Sec. 5.

3 Database and experimental setup

To evaluate the retrieval performance of the LPS descriptor without introducing noise in the results with approximate search approaches [10, 22], we chose a relatively small *UCID* database [20], on which we can perform an exact search. The performance of our LPS descriptors is compared to SIFT [11].

The whole *UCID* database contains 1338 images of size 512×384 pixels, divided into 262 unbalanced categories. After region detection and description, a single database entry for every category is constructed, comprising the descriptors from all the images of that category. Therefore, to equalize the database entry sizes as much as possible, different subsets of the *UCID* database were used in the experiments, where the number of examples per category is constant for each database subset (the required number of images is taken from larger categories in order provided by the ground truth). Tab. 1 summarizes the subsets of the database used for experiments presented herein.

A KD-Tree index [7] is then built based on the category descriptors, and stored for querying using the FLANN library [15]. We then perform a query with 1 image for every database category. The index performs a kNN search ($k = 7$) with each descriptor of a region detected on the query image. The final category is given through a voting mechanism where each nearest neighbor d_i of a query descriptor q_j will cast a vote for the category $cat(d_i)$ it belongs to:

$$vote(cat(d_i)) = \frac{100}{(L_1(d_i, q_j) + 0.1) \times |cat(d_i)|^{w_{cat}}}. \tag{5}$$

$L_1(d_i, q_j)$ refers to the distance between these two descriptors and $|cat(d_i)|$ is the number of descriptors in the category of the i-th nearest neighbor. Finally, w_{cat} is a parameter of the experimental setup. The five categories with the highest vote scores are examined in order to evaluate the performance of the descriptors.

The measures we used are mean average precision at five (MAP@5) and precision at one (P@1). Performance for different values of w_{cat} are shown in Fig. 1(a) and 2(d), but for all the summarized results, only the performance for the optimal w_{cat} value for each experiment is shown. This choice is made in order to present a fair comparison, and since not all the descriptors reach their peak performance for the same value of w_{cat}. This is additionally justified as this parameter is not present when using an approximate classification scheme.

4 Local pattern spectra

Local pattern spectra (LPS) are calculated from the selected MSER regions. As the two trees contain different regions, the descriptor for a maximal MSER will only be based on the max-tree, and similarly for the minimal MSERs.

The LPS are calculated like the global ones, except the calculation is done on the corresponding subtree. When calculating the LPS for the MSER region $\mathcal{L}^{k,i}$ in the tree, we only consider the attribute values of the descendants of the node. However, transitioning to the local version of the descriptor will introduce a new parameter influencing the scale invariance property of the descriptors.

Table 2: Parameters and their optimal values for the LPS.

symbol	significance	value SI-LPS	value SV-LPS
$m_{\mathcal{A}}$	upper bound for area	region size	
m_{NC}	upper bound for noncompactness	53	56
$N_b^{\mathcal{A}}$	number of area bins	9	10
N_b^{NC}	number of noncompactness bins	6	
M	scale parameter for the size attribute	20000	region size
$w(n_{1,1})$	normalized moment weights	20	
$w(n_{2,0}), w(n_{0,2}),$ $w(n_{4,0}), w(n_{0,4})$		10	

To achieve both the desired properties and competitive performance, the proposed descriptor is explained here through examining the experiments used to establish the best parameters. The summary of these parameters, explained individually henceforth, can be found in Tab. 2. Additionally, we consider combining the LPS with normalized central moments and enhancing the performance by adding the global pattern spectra. The influence of the database on the results is also discussed.

4.1 Scale invariance

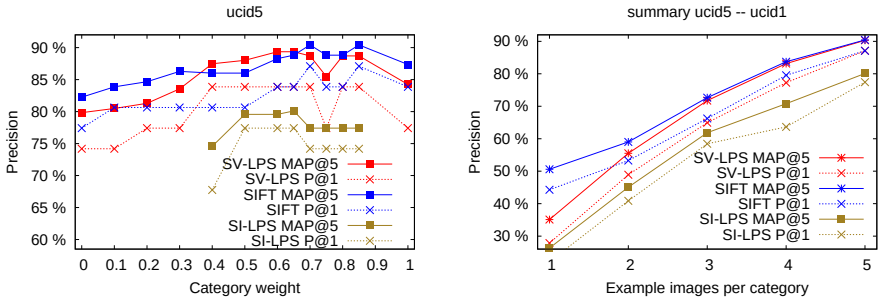
When calculating a global pattern spectrum for an entire image, the whole image size is used to determine the base of the logarithmic binning (especially if the database images are the same size [23,24]). If we choose to determine the binning base for each region separately based on the area of that region for the local descriptor scheme, the resulting LPS descriptor is not scale invariant.

Consider two version of the same region at different scales, with the area values belonging to the range $[1, m_1]$ and $[1, m_2]$ respectively. The scale invarianceproperty requires that, for a value $v_1 \in [1, m_1]$, the bin c_1 determined in the original scale is the same as the bin c_2 for the value $v_2 = v_1 \frac{m_2}{m_1}$ scaled to the range $[1, m_2]$. However, this is not the case for $m_1 \neq m_2$, as:

$$c_1 = \log_{\sqrt[N]{m_1}} v_1 \neq c_2 = \log_{\sqrt[N]{m_2}} v_2. \tag{6}$$

Therefore, to ensure the scale invariance of the descriptors, the area used to determine the binning and the logarithmic base have to be the same for all the regions. This area becomes then a parameter of the size attribute in LPS, called the *scale parameter* M.

Using a common scale M can be seen as rescaling all the regions to a reference scale, and has two consequences. First, for a region of size $m > M$, the minimal area value v of this region that can contribute to the spectrum when using a common binning is such that $v' = v\frac{M}{m} = 1$, meaning that all the (sub)regions with the area smaller then $\frac{m}{M}$ will be ignored. However, some particular regions



(a) (b)

Fig. 1: The results for the final version of the descriptors expressed in terms of mean average precision at 5 (MAP@5) and precision at 1 (P@1) for *ucid5* dataset for varying category weights are shown in (a). The results for *ucid5–ucid1* are summarized on (b) (performance shown for optimal weight w_{cat} for every dataset).

with a large enough area can still disappear when rescaling. This is the case for long thin objects with the width (along any dimension) small enough to downscale to under 1 pixel. Such regions should be ignored in the pattern spectrum, even if their attribute values fit with the binning. Because of this, we also determine the maximal possible value of the noncompactness attribute for all of the available area bins.

Second, the minimal area value (1 pixel) of a region of size $m < M$ will be rescaled to the value $v' = \frac{M}{m} > 1$, and the lower area bins at the common scale will be empty. The first area bin $c_{\min}$ that will contain information is then:

$$1 = b^{c_{\min}-1}\frac{m}{M} \rightarrow c_{\min} = \log_b \frac{M}{m} + 1. \tag{7}$$

We compare 2 versions of the descriptor: a) the *scale variant version* (SV-LPS), where the area of each region is used as the scale parameter M, and b) the *scale invariant version* (SI-LPS) where M is the same for all regions. The SV-LPS outperforms the SI-LPS on the performed experiments (cf. Fig. 1), and matches the SIFT performance. The best performance for the SI-LPS was obtained for $M = 20000$ (found experimentally) for *UCID* images. However, the UCID database is not very challenging in terms of scale change. We expect the SI-LPS performance to be less affected than that of SV-LPS when running experiments on a database focusing on scale change.

4.2 Binning parameters

With the area attribute, the upper bound used, m_A, is simply the size of the region: we can plausibly expect regions of all sizes lower than the size of the region itself to be present in its decomposition.

Examining the values of the noncompactness attribute for several images, we determined that very few regions have high values of this attribute. As such,

noncompactness values higher than a certain threshold can be safely ignored. Optimal values m_{NC} for both SV-LPS and SI-LPS were determined by examining the performance of the values close to the ones used in [23, 24]. Similar experiments were done to determine N_b^{NC} and N_b^A. The parameter tuning experiments are shown in Fig. 2.

The best parameters for the SI-LPS are easy to determine; we chose $N_b^{NC} = 6$ and $m_{NC} = 53$. For the number of area bins, we tested both $N_b^A = 8$ and $N_b^A = 10$ in the final descriptor combination (to be discussed in the following subsection). Using $N_b^A = 10$ produces better final results for SI-LPS, which are shown on Fig. 1. For SV-LPS, the influence of both parameters for noncompactness is much slighter. Surprisingly, we found that the optimal performance of SV-LPS reaches an optimum at the lower value of $N_b^A = 9$ (but a higher $m_{NC} = 56$ than SI-LPS). The optimal values for both SI-LPS and SV-LPS are listed in Tab. 2.

We also noted that using the optimal SI-LPS parameters in the scale variant version, we closely match the performance of our original parameter choice after the combination with image moments. Currently, no set of parameters is performing clearly better, but if future experiments confirm this behavior, it is still preferable to use a smaller N_b^A and decrease descriptor size.

4.3 Image moments and global pattern spectra

Five image moments, $n_{1,1}, n_{2,0}, n_{0,2}, n_{0,4}$ and $n_{4,0}$, were appended to a final version of all LPS descriptors. The weights resulting in the best performance (using the L_1 distance) were determined by examining the combination of the LPS and each of the moments separately. This weight is 20 for $n_{1,1}$ and 10 for other moments used. Additionally, an indicator value 2 is added to all the LPS descriptors originating from the max-tree, and 0 for the min-tree, thus additionally increasing the L_1 distance between any minimal and maximal MSERs.

Global pattern spectra on their own achieve MAP@5 around 70% on the *ucid5* dataset. They are added to the list of LPS for every image and treated equally to other local descriptors. The influence of combining these values with SI-LPS and SV-LPS for the optimal parameter choice is shown in Fig. 2(d).

4.4 Region size and database influence

Before calculating any descriptors in the evaluation framework of Mikolajczyk et al. [14], the region is first approximated by an ellipse with the same corresponding second moments, and then the region size is increased three times. Only then is the SIFT descriptor calculated using the provided implementation [11].

Since we want to be able to use the max-tree and the min-tree for the pattern spectra calculation, we chose to work with ancestor regions of the detected MSER such that the size of the parent is no larger than $xA(n_{k,i})$. We determined that, in order to get the same average area increase as in [14], we should use the value $x = 7.5$. The reason is that many regions have a much bigger parent region, which is then not considered, and the size increase is often smaller than x times.

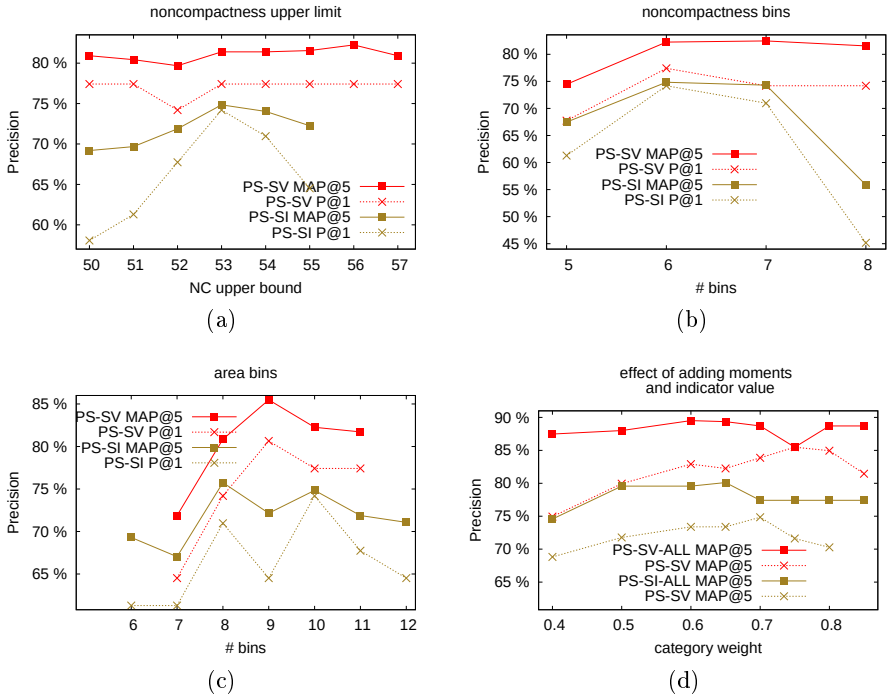


Fig. 2: Parameter tuning on *ucid5* database. The effect of varying the upper bound for noncompactness is shown on (a), similar for the amount of noncompactness bins on (b), and the area bins on (c). The effect of adding the moments and indicator value to the descriptor, with the best parameter settings is shown in (d). Note that the global descriptors for the SI-LPS are calculated with the scale value used for the other descriptors, and not using image size.

Fig. 1(b) summarizes the performance on all the subsets from Tab. 1, allowing us to examine the behavior of the descriptors for the increasing size of the database. The performance expectantly decreases with the increase of database size and decrease of the number of examples provided per category. As the separate influence of these two factors can not be determined just from experiments on these subsets, additional tests were carried out and analyzed in [4].

Besides the performance, it is important to note here that on the largest database subset used, the query speed for LPS is more than $4\times$ faster than that for SIFT (when the LPS descriptor of size 60 is used).

5 Remarks on the algorithm

The system was implemented in C++. The max-tree structure was used for both MSER detection and keypoint description. The non-recursive max-tree algorithm of [17] was used. This allows concurrent computation of the MSER

stability function (Eq. (1)), the area attribute and the moment of inertia, and the MSER. The method method is as follows:

- Compute the max-tree and min-tree according to [17].
- As the trees are built, compute:
  - local minima of the stability function, forming the sets of MSER regions,
  - attribute values for the nodes of the trees,
  - global pattern spectra [24].
- For each selected MSER region, repeat the computation of the pattern spectra locally in a sub-tree.
- Combine the attribute values, indicator value 0 or 2 and the pattern spectra to form a LPS descriptor for a MSER region.
- Add both global pattern spectra [23] corresponding to the whole image in the collection of descriptors for the image.

Unlike the calculation of global pattern spectra, the local pattern spectra use the constructed hierarchy but can not be computed concurrently because of different upper limits (for area) and binning scaling value.

However, in case of achieving improved results with the SI-LPS, adopting the scale invariant version to concurrent computation can be considered. While it would sacrifice true scale invariance, if the value M is used as a scale parameter, and we are calculating for a region of size m, we can set the largest bin to be $[b^{\lceil log_b m \rceil - 1}, b^{\lceil log_b m \rceil}]$, with the smallest bin having the upper bound $b^{\lceil log_b m \rceil - N_b}$. While it is then not always possible to get the values from the whole range of the largest bin, the bin values of the children can be used by their parents. When the upper bound of the largest bin changes, the child values can still be used with discarding the values from the smallest bin: the scale of those details is too low to be considered.

6 Discussion and conclusion

After successfully applying global pattern spectra in CBIR context [23, 25], we now attempt to construct a local region descriptor based on the pattern spectra. On the chosen subsets of the *UCID* database [20], the results obtained were better than when only using global pattern spectra (almost 20% in MAP@5 on *ucid5*), and matched the performance of the SIFT descriptor.

The proposed LPS descriptors have another advantage. In addition to the description calculation process being slightly faster for the pattern spectra than for the SIFT descriptors, our descriptors length is only 47% of the length of SIFT. This makes using these descriptors much faster – performing 262 queries on an index of the size 262 (*ucid1* dataset) took over 4 times longer using SIFT descriptors. This suggests that (especially in large scale CBIR systems), we can use more example images in order to enhance the precision, while still performing faster than SIFT.

As the performance of the descriptors depends on a lot of parameters, we need to explore a way to determine the optimal set of parameters automatically. Also,

while the LPS descriptors are rotation invariant, introducing scale invariance causes a decrease in performance. We plan to evaluate both the SI-LPS and SV-LPS on a database focused on scale changes to determine the value of true scale invariance in such cases.

Despite the parameters and the descriptor invariance which have to be further studied, matching the SIFT performance on the three subsets of the *ucid* dataset with a descriptor of less than half the length of SIFT is very promising. Additional successful experiments were performed and analyzed in [4]. It also prompts for evaluating the LPS performance with large scale CBIR system. It is probable that the results could be even further improved by combining the current LPS with pattern spectra based on other shape attributes, like in [23].

Lastly, the L_1 distance, designed to compare vectors of scalar values, is not the best choice for comparing histogram-like structures. Using different distances, or even divergences (e.g. [16]) which take into account the nature of the descriptor should also improve the performance.

References

1. Arandjelović, R., Zisserman, A.: Three things everyone should know to improve object retrieval. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. pp. 2911–2918. IEEE (2012)
2. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (SURF). Computer vision and image understanding 110(3), 346–359 (2008)
3. Benediktsson, J.A., Pesaresi, M., Arnason, K.: Classification and Feature Extraction for Remote Sensing Images from Urban Areas based on Morphological Transformations. IEEE Transactions on Geoscience and Remote Sensing 41(9), 1940–1949 (2003)
4. Bosilj, P., Kijak, E., Wilkinson, M.H.F., Lefèvre, S.: Short local descriptors from 2D connected pattern spectra, submitted to ICIP 2015[1]
5. Breen, E.J., Jones, R.: Attribute openings, thinnings, and granulometries. Computer Vision and Image Understanding 64(3), 377–389 (1996)
6. Dalla Mura, K., Benediktsson, J.A., Waske, B., Bruzzone, L.: Morphological Attribute Profiles for the Analysis of Very High Resolution Images. IEEE Transactions on Geoscience and Remote Sensing 48(10), 3747–3762 (2010)
7. Friedman, J.H., Bentley, J.L., Finkel, R.A.: An algorithm for finding best matches in logarithmic expected time. ACM Transactions on Mathematical Software (TOMS) 3(3), 209–226 (1977)
8. Hu, M.K.: Visual pattern recognition by moment invariants. Information Theory, IRE Transactions on 8(2), 179–187 (1962)
9. Jones, R.: Component trees for image filtering and segmentation. In: IEEE Workshop on Nonlinear Signal and Image Processing, E. Coyle, Ed., Mackinac Island (1997)
10. Lejsek, H., Jónsson, B.T., Amsaleg, L.: NV-Tree: Nearest Neighbors at the Billion Scale. In: Proceedings of the 1st ACM International Conference on Multimedia Retrieval. pp. 54:1–54:8. ICMR '11 (2011)

[1] preprint at `people.irisa.fr/Ewa.Kijak/preprint/icip.pdf`

11. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International journal of computer vision 60(2), 91–110 (2004)
12. Maragos, P.: Pattern spectrum and multiscale shape representation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 11(7), 701–716 (1989)
13. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide-baseline stereo from maximally stable extremal regions. Image and vision computing 22(10), 761–767 (2004)
14. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. Pattern Analysis and Machine Intelligence, IEEE Transactions on 27(10), 1615–1630 (2005)
15. Muja, M., Lowe, D.G.: Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. In: International Conference on Computer Vision Theory and Application VISSAPP'09). pp. 331–340. INSTICC Press (2009)
16. Mwebaze, E., Schneider, P., Schleif, F.M., Aduwo, J.R., Quinn, J.A., Haase, S., Villmann, T., Biehl, M.: Divergence-based classification in learning vector quantization. Neurocomputing 74(9), 1429–1435 (2011)
17. Nistér, D., Stewénius, H.: Linear time maximally stable extremal regions. In: Computer Vision–ECCV 2008, pp. 183–196. Springer (2008)
18. Ouzounis, G.K., Pesaresi, M., Soille, P.: Differential Area Profiles: Decomposition Properties and Efficient Computation. IEEE Transactions on Pattern Analysis and Machine Intelligence 34(8), 1533–1548 (2012)
19. Salembier, P., Oliveras, A., Garrido, L.: Antiextensive connected operators for image and sequence processing. Image Processing, IEEE Transactions on 7(4), 555–570 (1998)
20. Schaefer, G., Stich, M.: UCID: An Uncompressed Colour Image Database. In: Electronic Imaging 2004. pp. 472–480. International Society for Optics and Photonics (2003)
21. Schmid, C., Mohr, R.: Object recognition using local characterization and semi-local constraints. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(5), 530–534 (1997)
22. Sivic, J., Zisserman, A.: Video Google: Efficient visual search of videos. In: Ponce, J., Hebert, M., Schmid, C., Zisserman, A. (eds.) Toward Category-Level Object Recognition, LNCS, vol. 4170, pp. 127–144. Springer (2006)
23. Tushabe, F., Wilkinson, M.H.F.: Content-based image retrieval using combined 2D attribute pattern spectra. In: Advances in Multilingual and Multimodal Information Retrieval, pp. 554–561. Springer (2008)
24. Urbach, E.R., Roerdink, J.B.T.M., Wilkinson, M.H.F.: Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images. Pattern Analysis and Machine Intelligence, IEEE Transactions on 29(2), 272–285 (2007)
25. Urbach, E.R., Wilkinson, M.H.F.: Shape-only granulometries and grey-scale shape filters. In: Proc. Int. Symp. Math. Morphology (ISMM). vol. 2002, pp. 305–314 (2002)
26. Vedaldi, A., Fulkerson, B.: VLFeat: An Open and Portable Library of Computer Vision Algorithms. `http://www.vlfeat.org/` (2008)
27. Wilkinson, M.H.F.: Generalized pattern spectra sensitive to spatial information. In: Pattern Recognition, International Conference on. vol. 1, pp. 10021–10021. IEEE Computer Society (2002)