

Three-dimensional connectivity index for texture recognition[☆]



Joao B. Florindo^{a,b,*}, Gabriel Landini^c, Odemir M. Bruno^a

^aScientific Computing Group - São Carlos Institute of Physics, University of São Paulo, PO Box 369, 13560-970, São Carlos, SP, Brazil

^bInstitute of Mathematics, Statistics and Scientific Computing, University of Campinas, Rua Sérgio Buarque de Holanda, 651, Cidade Universitária "Zeferino Vaz" - Distr. Barão Geraldo, CEP 13083-859, Campinas, SP, Brasil

^cOral Pathology Unit, School of Dentistry, University of Birmingham, 5 Mill Pool Way, Edgbaston, Birmingham B5 7EG, United Kingdom

ARTICLE INFO

Article history:

Received 20 April 2016

Available online 29 September 2016

Keywords:

Local connectivity

Texture analysis

Pattern recognition

Image classification

ABSTRACT

This work proposes a new method of texture analysis for grey-level images based on the distribution of connectivity indexes in local neighbourhoods. The connectivity index acts as a measure of homogeneity of textures and its distribution is computed at various local neighbourhood sizes. The resulting descriptors provide an efficient multiscale representation of connectivity at different scales. The method was tested in the classification of UIUC, Outex, and KTH-TIPS2b databases and outperformed several state-of-the-art approaches, including such as LBP, LBP+VAR, MR8, multifractals among others.

© 2016 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The analysis of texture images has played a fundamental role in computer vision and pattern recognition during the last decades and the importance of this area of study is illustrated by the number of applications appearing in diverse areas such as Engineering [12], Medicine [1] and Physics [6].

Over the years, several approaches have been proposed to find methods capable of representing complex digital images using a reduced number of numerical features (or descriptors) that are relevant and non-redundant. Ideally, those features could be also useful for the analysis of objects and scenes in similar types of images.

Since the work of Haralick [8], considerable effort has been dedicated to find strategies that quantify how image pixels are distributed within local regions or neighbourhoods. Examples are the theory of textons [17], local patterns [14], local affine regions [11], invariants of scattering transforms [15], Fast Features Invariant to Rotation and Scale of Texture [16] and others [3]. Many of these efforts consider fixed sized neighbourhoods to extract basic parameters such as moments [8] and histograms [14,17]. Even though those approaches have demonstrated efficiency in both theory and practice, there might exist further methods which are more robust

and efficient in the increasingly challenging problems encountered when analysing digital images.

This work proposes using texture descriptors that combine the local features of pixel neighbourhoods with a multiscale analysis. The local arrangement of pixels is quantified by means of a connectivity index, similar to that proposed in [5] and which represents the number of connected pixels (in Graph Theory sense) to the central pixel of a neighbourhood. The degree of pixel clustering is estimated in a representation of the image where pixels are points in a quasi-three-dimensional space where x and y are the image coordinates and z is the pixel intensity function. The local measure is multiscale since it is computed over neighbourhoods (local windows) of various sizes. Finally, the descriptors are provided by the cumulative distribution of the connectivity index, as computed in [5], on the entire image and for the different neighbourhood sizes (radii).

Our proposed method has some particularities that distinguish it from previous textural analyses in the literature. First, unlike the method in [11] that focuses on particular affine regions, our method considers all pixels to have the same importance. Second, image invariants are not explicitly handled as in [11,15,16], although the symmetry of the connectivity index together with the multiscale approach used allows the quantification of such invariants and simplify the analysis where such invariants are not relevant, for example, if there is no significant variation or when objects at different scales or rotated should actually be considered as different instances. A third difference to other methods such as those in [8,14,17] is that despite the importance of pixel-level descriptions, our approach does not use directly the grey values

[☆] This paper has been recommended for acceptance by Y. Liu.

* Corresponding author. Fax: +55 16 3373 9879.

E-mail addresses: jbflorindo@ime.unicamp.br (J.B. Florindo),

G.Landini@bham.ac.uk (G. Landini), bruno@ifsc.usp.br (O.M. Bruno).

URL: <http://scg.ifsc.usp.br> (J.B. Florindo)

of the pixels or basic relations between neighbour pixels. Instead, a complex relation within the local neighbourhood is considered. This strategy takes into account how pixels are connected in the entire neighbourhood and this is based on real-valued distances instead of binary values at a particular distance.

The proposed method was assessed using the supervised classification of texture databases (a well-known benchmark problem of texture analysis). Three data sets commonly used in the literature were employed for this purpose, namely, UIUC [9], Outex [13], and KTH-TIPS2b [2]. To evaluate the classification performance, the proposed descriptors were also compared to other state-of-the-art and classical texture descriptors: Grey-Level Co-occurrence Matrix (GLCM) [8], Fourier [7], multifractals [19], Local Binary Patterns (LBP) [14], LBP+VAR [14], and MR8 [17].

The following section introduces the local connectivity index and other details of the proposed method. Section three discusses how the method relates to the concept of connectivity. Section four describes the validation experiments (databases and other compared descriptors), while section five shows and discusses the results obtained. Finally, section six briefly summarises the conclusions of the work.

2. Local connectivity index

The concept of local connectivity in image textures was first presented in [5] as a measure of local regularity in the context of a multifractal analysis. In simple terms, it is a generalisation to the three-dimensional Euclidean space of the local connected fractal dimension of binary images previously reported in [10,18].

In [10] a connectivity measurement is associated to each pixel p of the object (foreground) by counting the number of foreground pixels connected to p within a neighbourhood with radius r . The pixel p is “connected” to pixel $q \neq p$ means that there exists a set P of pixels satisfying

$$\begin{cases} p \in P \text{ and } q \in P \\ \forall (p_1, p_2) \in P : p_1 \neq p_2 \end{cases} \quad (1)$$

where $p_1 \neq p_2$ means that p_1 and p_2 are adjacent pixels (4 and 8-adjacency are classically utilised in binary images).

The ability of the connected fractal dimension to differentiate between images of normal and abnormal retinal vessels in [10] suggests that concept of connectivity might also be useful in other applications where the local regularity of the binary image needs to be assessed. Furthermore, the computation of the local descriptors for every position in the image also provides a complete and point-wise representation of the image data.

To extend the notion of connectivity to grey-level images, a strategy was proposed in [5] based on a pseudo-three-dimensional representation of grey-level images. Each pixel with coordinates (x, y) in the image I is mapped onto a point in quasi-three-dimensional Euclidean space E^3 with coordinates $(x, y, I(x, y))$.

Even though a similar definition of adjacency can be used in three-dimensional spaces (26-adjacency, for instance), this is not directly applicable to sparse sets of points such as those in the grey-level mapping. Therefore, a new definition for adjacency was established in [5] based on Euclidean distances between points: two points p with coordinates (x_1, y_1, z_1) and q with coordinates (x_2, y_2, z_2) , both in E^3 , are considered adjacent if they are at a distance smaller than a pre-defined threshold t :

$$p \neq q \Leftrightarrow \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \leq t. \quad (2)$$

The next step is to define the connectivity around a reference pixel p in the image. While in [5] a cube centred at p was used, here we use a ball or sphere to preserve the geometrical symmetry

in the local neighbourhood. In this way, a ball $B(p, r)$ centred at p with coordinates (x_p, y_p, z_p) with radius r is defined via

$$B(p, r) = \{(x, y, z) : \sqrt{(x - x_p)^2 + (y - y_p)^2 + (z - z_p)^2} \leq r\}. \quad (3)$$

The connectivity index can then be informally defined as the number of points inside $B(p, r)$ and connected to p . The connected set therefore corresponds to the set of points within B , and including p , such that each point q_a has at least one other point q_b at a distance smaller than t . It is important to point out that pixels connected to p but not belonging to $B(p, r)$ (at a distance from p larger than r) are not counted in the composition of the connected set. The connectivity index $C(p, r, t)$ can be summarised as

$$C(p, r, t) = \#\{p\} \cup \{q_a : q_a \in B(p, r) \text{ and } \exists q_b \in B(p, r) \text{ and } D(q_a, q_b) < t\}, \quad (4)$$

where D is the Euclidean distance, \cup represents the classical notation for set union or, equivalently in this case, concatenation and $\#$ symbolizes the cardinality (number of elements) of the set.

A flowchart in Fig. 1 shows the basic procedure to compute the connectivity of a particular pixel p_0 in the image. More details and a pseudo-code is provided in [5] for the interested reader.

3. Proposed method

From its definition, the connectivity index presented earlier can be computed locally, relative to each pixel in the image. Thus, this corresponds to a transform of the original image where each pixel is converted into an integer non-negative value representing the local connectivity. Same as in numerous other transforms (Fourier, wavelets, etc.) these mappings constitute powerful tools to provide a new viewpoint on the image and therefore might be useful to describe various patterns that cannot be discriminated in the original space representation.

Fig. 2 illustrates the local connectivity for four images from the UIUC database [9], belonging to two different classes. Two major features can be noted in this representation. Firstly, the transform conveys information which is different from that represented by the grey levels while still preserves some prominent characteristics (for instance, the cracked edge in the left figure). Secondly, images from diverse classes lead to dissimilar patterns in the connectivity image version.

Another interesting property of the connectivity measure is represented in Fig. 2 where the connectivity indexes were computed for different radii r (3, 4 and 5 pixels) and shown as a colour map: apart from the increased magnitude of the indexes (which is expected, provided that larger neighbourhoods contain more points) different patterns emerge in the distribution of the connectivity indexes at different radii. This suggests the multiscale nature of the connectivity, making it potentially useful to describe complex images. In particular, the ability to distinguish the classes presented in Fig. 2 is boosted by considering the image transform results at different radii.

The distinction of classes appears more evident in Fig. 3 resulting from different distributions across classes at different sized neighbourhoods. Such example illustrates that whereas the connectivity index shows its importance in a multifractal context in [5], it also could be a useful local descriptor on its own. While the performance of the connectivity is illustrated here on a small subset of images from UIUC database, the Results section presents a more robust experiment over the whole database, confirming the power of our approach.

Based on these observations, we propose a method for computing descriptors based on the combination of local connectivity transform histograms for various radii r and thresholds t .

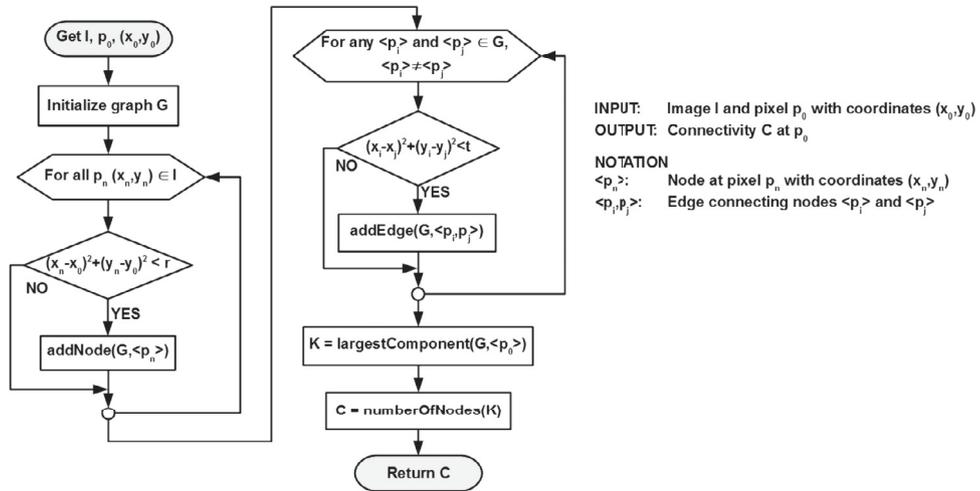


Fig. 1. Flowchart used to compute the connectivity of a pixel in the image.

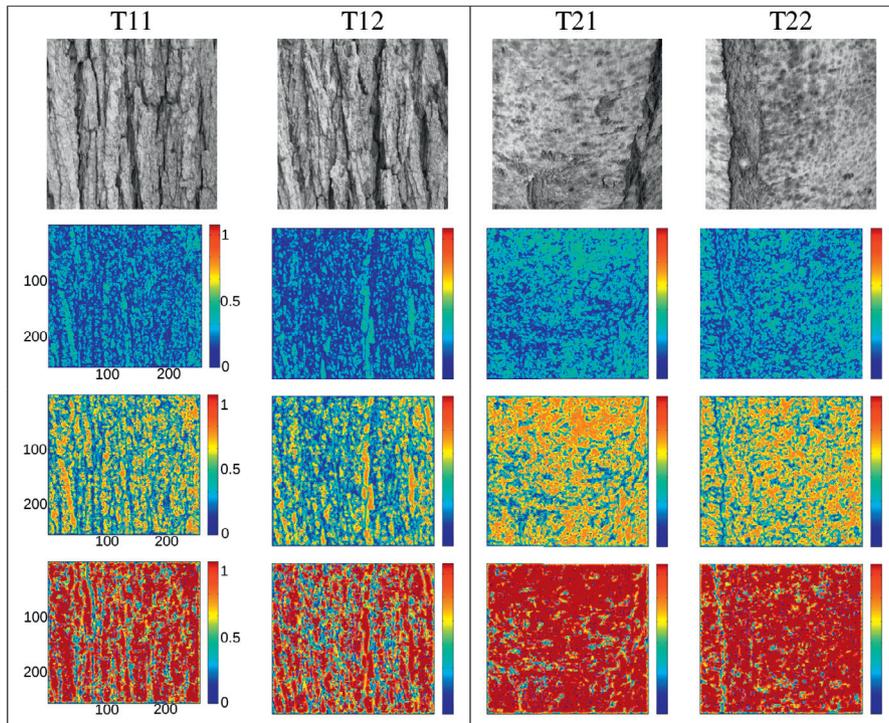


Fig. 2. Connectivity indexes in a colour map (normalised between 0 and 1) with different values of radius (from top to bottom, $r = 3$, $r = 4$ and $r = 5$) for four UIUC images, from two distinct classes.

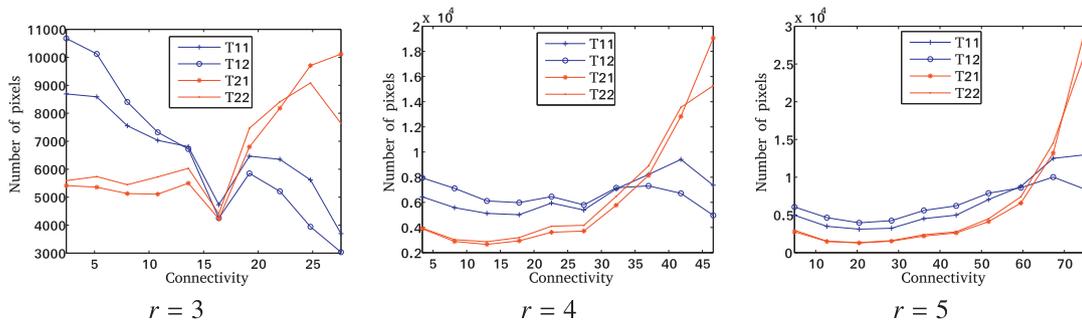


Fig. 3. Histograms distinguishing images in Fig. 2 for different values of radius ($r = 3$, $r = 4$ and $r = 5$). T11, T12, T21 and T22 are the respective texture images in Fig. 2.

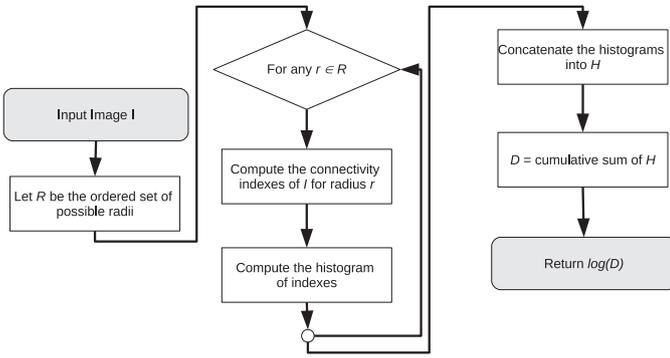


Fig. 4. A flowchart of the proposed method.

While radii between 1 and r_{max} pixels are considered, the possible values of r are those that correspond to the underlying discrete Euclidean distances, expressed by the set \mathfrak{R} in the following expression:

$$\mathfrak{R} = \{r \in \mathbb{N} : \exists(x, y, z) \in \mathbb{Z}^3 \text{ satisfying } \sqrt{x^2 + y^2 + z^2} = r\}. \quad (5)$$

Let us assume that $\mathfrak{C}(r, t)$ is the increasingly ordered set of unique values of $C(p, r, t)$ computed for any $p \in I$. For each radius $r \in \mathfrak{R}$ the cumulative histogram h_r^t is defined for each counter index $k \in \mathfrak{C}$ by

$$h_r^t(k) = \sum_{p \in I, k' \leq k} \delta(C(p, r, t) - k'), \quad (6)$$

where $\delta(x)$ stands for the Dirac delta function (one at $x = 0$ and zero everywhere else).

Next, the sets of features in h_r^t for all radii are concatenated into a feature vector \mathfrak{H}^t :

$$\mathfrak{H}^t = \bigcup_{r \in \mathfrak{R}} h_r^t. \quad (7)$$

Finally, the proposed descriptors \mathfrak{D}^t for image I are obtained from the cumulative sum of \mathfrak{H}^t followed by a logarithm operation:

$$\mathfrak{D}^t(u) = \log \sum_u \mathfrak{H}^t \leq u. \quad (8)$$

There are two reasons to use the cumulative sum instead of an L^2 -histogram: first, it ensures more regular features because the

sum provides a monotonically increasing function; second, its values increase according to the variation in the original data (histogram). These two characteristics are important in the distance-based classifiers used here because the dissimilarity metric is more sensitive to the accumulated variation. The logarithm can be justified in a similar way: the connectivity index scales exponentially as described in [5] and the logarithm attenuates the wide range of scales in the feature space. Note that more than one value of t could be used. After the descriptors are computed, the resulting features are concatenated, and to reduce the dimensionality a large number of features, a Karhunen–Loève transform is applied to preserve only the most relevant and uncorrelated descriptors. Fig. 4 summarizes the steps involved in the method in a flowchart, while Fig. 5 illustrates each step with a visual example.

4. Experiments

The proposed descriptors were assessed in the classification of three established databases of texture images, namely, UIUC [9], Outex [13], and KTH-TIPS2b [2]. UIUC is composed of 1000 grey-level texture images with dimension 256×256 equally divided into 40 groups, containing photographs taken under uncontrolled conditions and with relevant variation in scale and viewpoint. The version of Outex database used here is the suite Outex_TC_00013 in [13] and contains 1360 colour images (here converted to a grey-scale) captured under controlled conditions of illumination and imaging geometry. The samples are divided into 68 groups (20 images with dimension 128×128 in each one). Finally, KTH-TIPS2b is a set of material images, collected under varied conditions of illumination, scale and pose, comprising 11 materials (classes), each one with 4 samples and each sample with 108 cropped images with dimension of 200×200 pixels.

The connectivity features were compared to other state-of-the-art and classical approaches for texture classification in the literature, i.e. Grey-Level Co-occurrence Matrix (GLCM) [8], Fourier [7], multifractals [19], Local Binary Patterns (LBP) [14], LBP+VAR [14], and MR8 [17].

The classification of the databases according to the compared descriptors was carried out using a linear discriminant classifier (LDA) [4] after a principal component analysis (PCA) [4] to reduce the correlation among features in all compared approaches. Such approach for the classification prevents redundant information from being considered in the segmentation of the feature

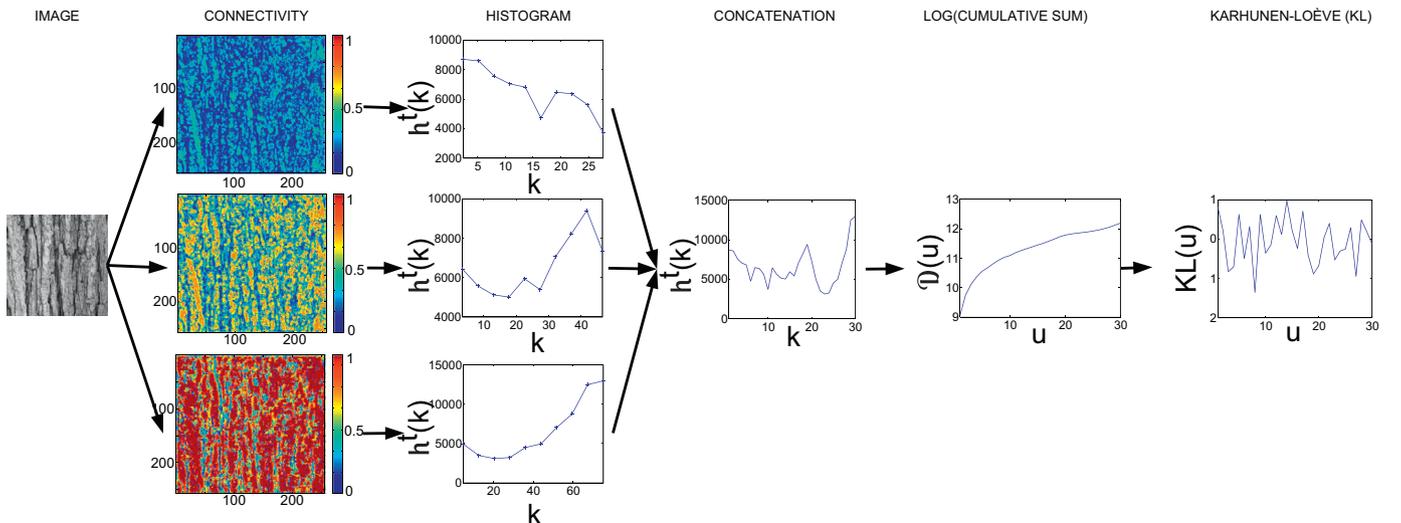


Fig. 5. Steps involved in the proposed method. From left to right, the original image, the connectivity for different radii, connectivity histograms, concatenated histograms, cumulative sum and Karhunen–Loève transform.

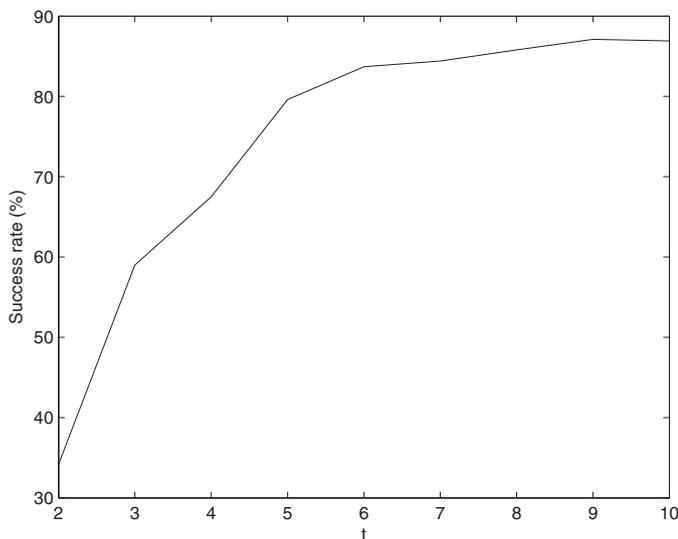


Fig. 6. Percentage of UIUC images correctly classified when the value of t is varied between 2 and 10.

Table 1
Best success rates for the compared descriptors.

Method	UIUC	Outex	KTH-TIPS2b
Fourier	64.00 ± 0.03	82.21 ± 0.02	72.50 ± 0.01
GLCM	58.70 ± 0.03	68.23 ± 0.05	65.05 ± 0.01
Multifractals	82.40 ± 0.03	75.07 ± 0.03	69.95 ± 0.01
LBP	77.40 ± 0.05	74.63 ± 0.02	70.16 ± 0.01
LBP+VAR	83.90 ± 0.02	76.03 ± 0.04	70.73 ± 0.02
MR8	84.70 ± 0.02	71.54 ± 0.04	73.63 ± 0.02
Proposed	88.60 ± 0.03	84.26 ± 0.03	83.54 ± 0.01

space. The simultaneous dimensionality reduction and decorrelation is achieved by a combination of PCA with the canonical analysis of LDA, which takes into account the distribution of features among the classes in the training set. These characteristics make the adopted classifier more appropriate for the proposed descriptors, which are naturally highly correlated as a consequence of the iterative process they follow (descriptors from neighbourhoods of larger radii enclose part of the information expressed by those of smaller radii). A maximum of 100 PCA scores were used in the LDA classification. Considering that we were working on databases where the classes are known in advance, the classification was considered to be supervised. A division into training and testing sets following a 10-fold scheme was applied for cross-validation.

The parameter r_{max} was set to 5 in all tests based on the observation in [5] that larger radii severely increase the computational cost and number of descriptors without a significant gain in the classification performance. On the other hand, the values for the parameter t were empirically obtained from a training set of images. Fig. 6 illustrates the variation in the rate of images correctly classified in UIUC database when t ranges between 2 and 10. In the end, $t = 10$ was used for UIUC, a combination of $t = 2, 3, 4, 5$, and 7 constituted the optimal parameters for Outex, and a combination of $t = 5, 6, 8$, and 9 was used to classify KTH-TIPS2b materials. As mentioned earlier, all the connectivity indexes were concatenated before the application of PCA to reduce dimensionality.

5. Results

Table 1 lists the rates of correct classification (i.e. the ratio of images correctly classified) for each database and each descriptor in comparison with the proposed method. The table suggests that classifying the Outex and KTH-TIPS2b databases was some-

what more challenging than the UIUC images, as illustrated by the smaller correctness rates and higher cross-validation errors. This behaviour is expected because of the nature of each database, because the high number of classes in Outex and samples in KTH-TIPS2b data set. Such complexity is the rationale, for example, for the need of including more than one value of t to generate the feature vector. In all cases, the connectivity descriptors outperformed all the other approaches by at least 5% (UIUC), 2% (Outex), and 13% (KTH-TIPS2b) in relative percentages. In KTH-TIPS2b the classification performance was also better than that reported in [16] (76%) using a similar database and protocol.

Beyond the correctness rate, a more complete and precise metric to assess the performance of supervised classification is the confusion matrix. This expresses the correctness in each class as well as the number of samples misclassified, including which pairs of classes A/B introduced confusion (e.g. samples from class A but assigned to B by the classifier and vice-versa). Fig. 7 shows the confusion map for UIUC, Outex, and KTH-TIPS2b, when classified by the best two methods in Table 1. The colour bars represent the number of images assigned to each particular class. Although visually the differences may not appear evident, the advantage of the connected descriptors is expressed in the colour map by a diagonal with higher values (i.e. more red points). Less confusion is also observed in the most critical regions of the map, for instance, in the last 10 classes (right end of the plot) of the Outex dataset or in the lack of yellow squares for the connected descriptors in KTH-TIPS2b. In practice, the information enclosed by this type of representation can be used to identify classes of materials with high potential of being misclassified. For example, in KTH-TIPS2b, classes 9 and 8 are confused to a significant extent by MR8 but not by the proposed descriptors. The corresponding materials (white bread and linen) are examples of textures that are more appropriately handled by a multiscale approach as that employed by the connected descriptors.

The results confirm the efficiency of the connectivity descriptors as a means of expressing the richness and complexity encompassed by the pixel values of a grey-level image in a reduced and more easily manageable set of numerical values. The classification performance can be justified by the ability of the connectivity index to quantify certain pixel relations within a local neighbourhood. Varying the size of the neighbourhoods makes the analysis multiscale and this helps describe the heterogeneity of the texture. This is important information for both natural and artificial vision systems.

6. Conclusions

This work presented a new method, named connectivity descriptors, to extract features from texture grey-scale images. The approach is based on the computation of a connectivity index within a local image neighbourhood. This corresponds to the number of pixels more closely related to the central pixel in the neighbourhood. The method performance was compared to other methods in the literature and tested on three well established texture databases.

The connectivity descriptors provided the best classification rates among the comparison of methods for both database tests. Such promising results can be supported by the ability of connectivity indexes in quantifying the relationships among pixels within a local region of the image. This, added to the multiscale nature of the method, makes the features sufficiently robust and precise to distinguish between images of different textures even in the complex scenarios of the datasets investigated.

Finally, given the generic nature of the UIUC, Outex, and KTH-TIPS2b databases, the results suggest that the connectivity descriptors might be suitable for a number of real-world applications in

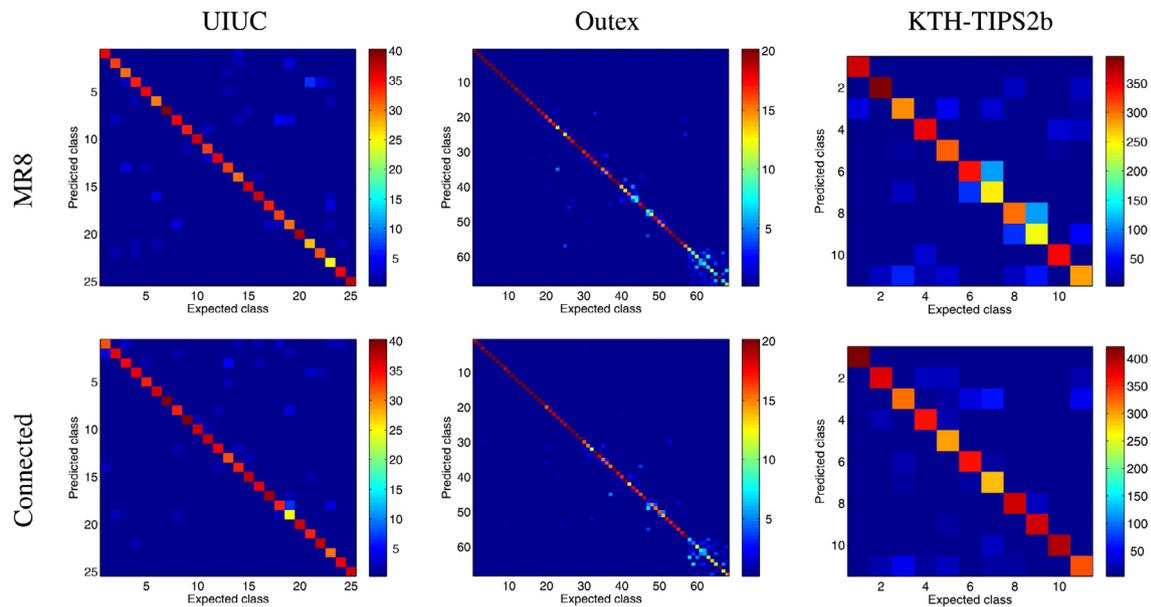


Fig. 7. Confusion matrices for the best two methods in Table 1 on UIUC, Outex, and KTH-TIPS2b. The scale bar represents the number of images pertaining to an expected class and assigned by the classifier to a predicted class.

problems involving the recognition and quantification of patterns in digital images.

Acknowledgements

J. B. Florindo gratefully acknowledges the financial support of FAPESP (The State of São Paulo Research Foundation) Grant #2012/19143-3 and #2013/22205-3. G. Landini acknowledges support by the Engineering and Physical Sciences Research Council (EPSRC), UK, through funding under grant EP/M023869/1 Novel context-based segmentation algorithms for intelligent microscopy. O. M. Bruno gratefully acknowledges the financial support of CNPq (National Council for Scientific and Technological Development, Brazil) (Grant #307797/2014-7 and Grant #484312/2013-8) and FAPESP (Grant #11/01523-1).

References

- [1] C. Arteta, V. Lempitsky, J.A. Noble, A. Zisserman, Detecting overlapping instances in microscopy images using extremal region trees, *Med. Image Anal.* 27 (2016) 3–16.
- [2] B. Caputo, E. Hayman, P. Mallikarjuna, Class-specific material categorisation, in: *ICCV*, IEEE Computer Society, 2005, pp. 1597–1604.
- [3] M. Cimpoi, S. Maji, A. Vedaldi, Deep filter banks for texture recognition and segmentation, in: *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [4] R.O. Duda, P.E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.
- [5] J.B. Florindo, G. Landini, O.M. Bruno, Texture descriptors by a fractal analysis of three-dimensional local coarseness, *Digital Signal Process.* 42 (2015) 70–79.
- [6] Y. Gao, M.E. Helgeson, Texture analysis microscopy: quantifying structure in low-fidelity images of dense fluids, *Opt. Express* 22 (2014) 10046–10063.
- [7] R.C. Gonzalez, R.E. Woods, *Digital Image Processing*, (2nd ed), Prentice Hall, Upper Saddle River, N.J., 2002.
- [8] R.M. Haralick, Statistical and structural approaches to texture, *Proc. IEEE* 67 (1979) 786–804.
- [9] K. Kilic, R. Abiyev, Exploiting the synergy between fractal dimension and lacunarity for improved texture recognition, *Signal Process.* 91 (2011) 2332–2344.
- [10] G. Landini, P.I. Murray, G.P. Misson, Local connected fractal dimensions and lacunarity analyses of 60 degrees fluorescein angiograms, *Invest. Ophthalmol. Visual Sci.* 36 (1995) 2749.
- [11] S. Lazebnik, C. Schmid, J. Ponce, A sparse texture representation using local affine regions, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (2005) 1265–1278.
- [12] Y.T. Luo, L.Y. Zhao, B. Zhang, W. Jia, F. Xue, J.T. Lu, Y.H. Zhu, B.Q. Xu, Local line directional pattern for palmprint recognition, *Pattern Recognit.* 50 (2016) 26–44.
- [13] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen, S. Huovinen, Outex - new framework for empirical evaluation of texture analysis algorithms, in: *ICPR*, 2002a, pp. 701–706.
- [14] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (2002b) 971–987.
- [15] L. Sifre, S. Mallat, Rotation, scaling and deformation invariant scattering for texture discrimination, in: *CVPR*, IEEE Computer Society, 2013, pp. 1233–1240.
- [16] M. Sulc, J. Matas, *Fast Features Invariant to Rotation and Scale of Texture*, Springer International Publishing, Cham, 2015, pp. 47–62.
- [17] M. Varma, A. Zisserman, A statistical approach to texture classification from single images, *Int. J. Comput. Vision* 62 (2005) 61–81.
- [18] R.F. Voss, J.C. Wyatt, Multifractals and the local connected fractal dimension, in: A. Crilly, R. Earnshaw, H. Jones (Eds.), *Applications of Fractals and Chaos*, Springer Berlin Heidelberg, 1993, pp. 171–192.
- [19] Y. Xu, H. Ji, C. Fermüller, Viewpoint invariant texture description using fractal analysis, *Int. J. Comput. Vision* 83 (2009) 85–100.