

Locally Invariant Fractal Features for Statistical Texture Classification

Manik Varma
Microsoft Research India
manik@microsoft.com

Rahul Garg
Indian Institute of Technology Delhi
rahul.gargrahul@gmail.com

Abstract

We address the problem of developing discriminative, yet invariant, features for texture classification. Texture variations due to changes in scale are amongst the hardest to handle. One of the most successful methods of dealing with such variations is based on choosing interest points and selecting their characteristic scales [Lazebnik et al. PAMI 2005]. However, selecting a characteristic scale can be unstable for many textures. Furthermore, the reliance on an interest point detector and the inability to evaluate features densely can be serious limitations.

Fractals present a mathematically well founded alternative to dealing with the problem of scale. However, they have not become popular as texture features due to their lack of discriminative power. This is primarily because: (a) fractal based classification methods have avoided statistical characterisations of textures (which is essential for accurate analysis) by using global features; and (b) fractal dimension features are unable to distinguish between key texture primitives such as edges, corners and uniform regions.

In this paper, we overcome these drawbacks and develop local fractal features that are evaluated densely. The features are robust as they do not depend on choosing interest points or characteristic scales. Furthermore, it is shown that the local fractal dimension is invariant to local bi-Lipschitz transformations whereas its extension is able to correctly distinguish between fundamental texture primitives. Textures are characterised statistically by modelling the full joint PDF of these features. This allows us to develop a texture classification framework which is discriminative, robust and achieves state-of-the-art performance as compared to affine invariant and fractal based methods.

1. Introduction

In this paper, we address the problem of classifying single images of textures obtained under unknown viewpoint and illumination conditions. As is well documented by now, this is an extremely demanding task, specially under such unconstrained settings, due to the large intra-class and small



Figure 1. Changing scale can have a dramatic impact on the appearance of a material: (a) and (b) are examples of leaf and (c) and (d) of skin taken at different scales while keeping all other parameters constant.

inter-class variation that textures exhibit. It therefore becomes crucial to develop texture features that are not only discriminative across many classes but also invariant to key transformations, such as rotation, scaling, affine illumination changes, etc.

Scale variations can have a dramatic impact on the imaged appearance of a texture (see Figure 1). While no locally invariant method can handle scale changes of such magnitude, our goal in this paper is to build descriptors invariant to scale changes in the range shown in Figure 6.

Scale variations in textures are amongst the hardest to handle and only modest progress has been made in coming up with scale invariant features [6, 7, 14, 25, 26, 39]. The most promising methods which have demonstrated good performance on real world datasets are [20, 37, 38]. The approach followed in [20, 38] is based on the affine adaptation process. First, certain interest points are chosen and then their characteristic scale determined by selecting a local region for which the Laplacian operator achieves a maximum. Further processing is done and an affine viewpoint and illumination invariant descriptor is computed based on the selected region. The technique in [37] is based on a fractal approach which computes a global multi-fractal spectrum (MFS) vector formed over three different measures.

While features derived from interest points and the affine adaptation process have proved to be very useful for matching, they could be inappropriate for many texture classification tasks. Firstly, interest point detectors typically produce a sparse output and might miss important texture prim-

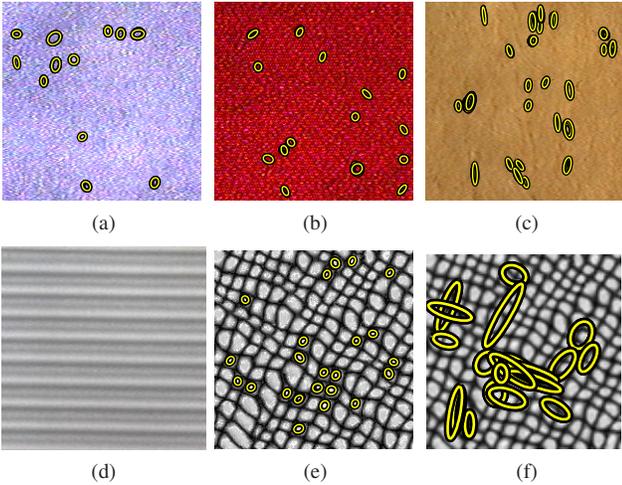


Figure 2. The affine adaptation process might not be suitable for textures: (a)-(c) are images of felt, velvet and lambswool that do not possess characteristic scales. Hence, arbitrary regions are detected; (d) is an image of paper where no interest points were found; (e) is an image of a Brodatz texture; and (f) shows the detected regions when (e) is scaled down by half and then rescaled back (after region detection) to original size for comparison purposes. Note that the regions in (e) and (f) do not correspond. All regions were detected using the Hessian-Affine detector [23].

itives. Secondly, a sparse output in a small image might not produce enough regions for a robust statistical characterisation of the texture. As regards the affine adaptation process, there exist many classes of textures which do not possess a characteristic scale. Furthermore, there are always issues regarding the repeatability of the detector and the stability of the selected region [23]. As a result, features computed via such a process can be unreliable (see Figure 2).

Fractals present a mathematically well founded alternative to dealing with the problem of scale. They have found application in classification [4, 10, 15, 24, 34, 37], segmentation [5, 13, 17, 36], synthesis [16] and other important texture problems [1, 19, 22, 33, 35]. However, the performance of fractal based texture classifiers has often lagged behind the state-of-the-art. This is primarily due to two reasons.

Firstly, fractal based classification methods have traditionally avoided a statistical characterisation of textures preferring to model them using globally computed fractal dimensions or MFS vectors. This is a significant shortcoming as textures have often shown to be best described by the statistical distribution of textons [3, 8, 12, 21, 27, 31]. For instance, to make an analogy with filter banks, considerable progress was made in both classification and synthesis by modelling textures using first the mean, then the mean and variance and finally by the full joint PDF of locally computed filter responses [30]. As such, it should be expected that the same progression should yield superior results in

the case of fractal features as well. Note that local fractal features are already the norm in fractal based texture segmentation, where many methods [5, 13, 17, 36] proceed by first calculating local fractal features, then clustering and labelling them followed by various post-processing steps such as boundary smoothing.

Secondly, many point sets such as those representing image corners, edges, homogeneous regions and other important texture primitives have identical fractal dimensions. As a result, the performance of fractal based classification schemes has often not been as good as other competing methods. For instance, classification results obtained by the leading fractal based method [37] are inferior to those obtained by affine adaptation [20] on the UIUC dataset.

In this paper, we overcome these drawbacks and develop local, fractal based features which can be evaluated densely. The features are robust as they do not depend on characteristic scale selection or the affine adaptation process. Furthermore, it is shown that the local fractal dimension is invariant to local bi-Lipschitz transformations, such as local affine or perspective projections and certain smooth, non-linear transformations. In addition, we develop a new fractal feature for texture classification, the local fractal length, and show that it can distinguish between important texture primitives such as edges, corners, uniform regions, etc. at the cost of reduced invariance. Textures are then characterised statistically by modelling the full joint PDF of these features. This allows us to develop a texture classification framework which is discriminative, robust and achieves good performance as compared to the state-of-the-art affine invariant and fractal based methods [20, 37].

The rest of the paper is organised as follows. Section 2 briefly reviews applicable fractal theory and how it has traditionally been used for texture classification and segmentation. Next, local fractal features are developed and their discriminative power and invariance properties discussed. In Section 3, the new features are empirically validated on the UIUC [20] and CURET [9] databases. We conclude in Section 4 and explore avenues of future work.

2. Fractal Features

Before developing the proposed local fractal features and discussing their properties, we briefly review fractal theory as applicable in our scenario.

Review Very loosely speaking, a perfect fractal is a shape which, amongst other things, appears similar at all scales of magnification. Due to this property, a perfect fractal can be decomposed into N similar copies of itself, each scaled down by a factor s , which tile the original shape exactly. It often turns out to be the case that the quantities N and s are related by a power law, i.e. $N(s) \propto s^{-D}$ where D is defined

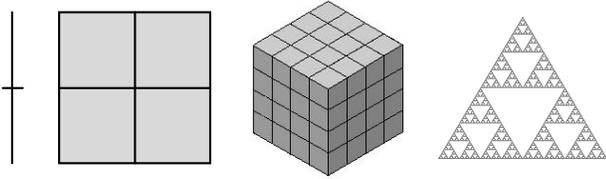


Figure 3. Each of the shapes can be decomposed into N similar copies of itself scaled by a factor of $s = 0.5$. The number of half-scaled copies N equals 2, 4 and 8 for the line, square and cube respectively. This leads to fractal dimensions of $D = -\log N / \log s = 1, 2$ and 3 as expected. However, the Sierpinski triangle, by construction, has only $N = 3$ half-scaled copies of itself leading to a non-integral fractal dimension $D = -\log 3 / \log 0.5 = 1.585$.

to be the fractal dimension of the shape (though there exist many other definitions too [11]). For smooth shapes from classical geometry, such as lines, squares, cubes, etc. the fractal dimension equals the topological dimension. However, for irregular point sets, the fractal dimension is not an integer but lies between the bounding topological dimensions (see Figure 3). As such, it can be loosely interpreted as a factor governing the irregularity of the point set or sometimes as the roughness of the shape. The concept of fractal dimension has been generalised to many cases where the point set being considered is not a perfect fractal [11], such as multi-fractals and statistically self similar fractals.

There are three primary ways in which fractal features have been computed from images and applied in texture analysis. In the first, the surface which generated the image is modelled as a fractal, typically using a fractional Brownian motion model, and its roughness calculated in terms of the Hurst parameter or extensions [10, 15]. The surface roughness then acts as a parameter for discrimination between classes. In the second method, an image I is directly modelled as an intensity surface $(x, y, I(x, y))$ and its fractal dimension is used to parameterise the texture [17, 24, 36]. Finally, an image can also be seen as a union of point sets, each of whose fractal dimension is taken together to form an MFS vector [34, 37]. However, in much of previous research, the focus has been on a scale variant analysis of textures rather than deriving scale invariant features.

The method of Xu *et al.* [37] falls in the third category. In it, a 26 dimensional global MFS vector is computed per measure. Given a texture image, its q -th measure moment is calculated as a sum over a partitioning of the image into non-overlapping boxes of length r . Assuming that the moment varies with r as r^β , the MFS vector is calculated from β via a Legendre transform. Three measures are used in all (Gaussian, energy and Laplacian) and the individual MFS vectors concatenated to give a final 78 dimensional MFS feature vector. In essence, the MFS vector is a collection

of fractal dimensions and therefore avoids a statistical description of the texture. The MFS vector is proved to be invariant to geometric global bi-Lipschitz transforms and to multiplicative changes in the illuminant intensity (full affine illumination invariance can't be achieved due to properties of the Gaussian measure). The MFS vector is quick to compute and achieves good results on the UIUC database using nearest neighbour classifiers. Nevertheless the performance is inferior to that of [20]. The authors note that this is primarily due to the MFS vector being relatively less robust to illumination changes and also as the images in the UIUC database are not large enough for stable MFS vector computation. However, as will be shown, the poor performance is also due to the use of globally computed fractal dimensions which are not very discriminative.

Local fractal features Our approach is based on the assumption that, given a suitable *measure* μ , the “size” of local point sets in textured images follows a local power law. To take a concrete example, given an image I , let $\mu(\bar{B}(\mathbf{x}, r))$ be the sum of all pixel intensities that lie within a closed disk \bar{B} of radius r centred at an image point \mathbf{x} , i.e. $\mu(\bar{B}(\mathbf{x}, r)) = \sum_{\|y-\mathbf{x}\| \leq r} I(y)$. We hypothesise that

$$\mu(\bar{B}(\mathbf{x}, r)) \propto r^{D(\mathbf{x})} \quad (1)$$

$$\Rightarrow \log \mu(\bar{B}(\mathbf{x}, r)) = D(\mathbf{x}) \log r + L(\mathbf{x}) \quad (2)$$

where $D(\mathbf{x})$ is the local fractal dimension, also known as the Hölder exponent. While this power law assumption might appear overly restrictive at first, it turns out to be a surprisingly good approximation for many real world cases [11]. In particular, Figure 4 illustrates the quality of the approximation on four real world texture images taken from the UIUC database. The $\log \mu$ versus $\log r$ plots for eight image points are shown in the graph. As can be seen in each of the eight cases, the points lie along a straight line indicating that the power law is being followed faithfully.

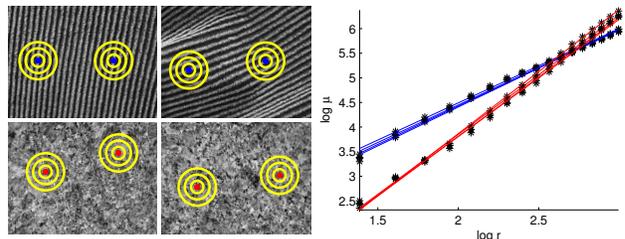


Figure 4. The $\log \mu$ versus $\log r$ plots for 8 points in 4 images from the UIUC texture database are shown. Firstly, all the 8 plots are straight lines indicating that the local power law assumption holds true. Secondly, all 4 lines from the same class are clustered together (in fact, for each r , the 4 points per class appear near coincident) with a distinct difference between the 2 classes. This shows that local fractal features can be used to distinguish textures.

While there exist many methods for estimating the fractal dimension, we choose to simply read off $D(\mathbf{x})$ and $L(\mathbf{x})$ as the slope and intercept of the $\log \mu$ versus $\log r$ graph respectively (more sophisticated techniques can only improve estimates). It should be noted that estimates of the slope and intercept can be obtained densely, without relying on any pre-defined interest point detector. Furthermore, no specific characteristic scale is chosen. In fact, robust line fitting methods can be utilised so that even if there are noisy measurements at a few scales (which could influence characteristic scale selection) the slope and intercept of the fitted line remain relatively stable.

Discriminative power and invariance While it might be plausible to apply the fractal model to textures, we still have to determine how useful fractal features are for classification. In fact, it turns out that the local fractal dimension by itself is not very discriminative. As Figure 5 demonstrates, many fundamental texture primitives, such as homogeneous regions, edges, corners, etc. have identical local fractal dimension $D(\mathbf{x})$. However, the local intercept $L(\mathbf{x})$ is a good feature for distinguishing between such primitives. In general, $\exp(L)$ can be interpreted as the D -dimensional fractal length of μ , i.e. the “size” of μ when measured with a unit of “size” r^D . As such, we define $L(\mathbf{x})$ to be the local (log) fractal length. While the idea of lacunarity [17, 32] has been experimented with to distinguish fractals with identical dimension, fractal lengths have not been used as features for texture classification to the best of our knowledge.

The fact that local fractal features can be used to distinguish texture classes can also be seen from Figure 4. Two images of fabric and two of granite from the UIUC texture database are shown along with $\log \mu$ versus $\log r$ plots of eight image points. The lines from the fabric points are clustered together as are the lines from the granite points, though the two sets are distinct.

As regards invariance, any image transformation which leaves the set of μ s unchanged, such as image rotation, will have no impact on D and L (affine illumination transformations can also be incorporated in this category). Furthermore, for perfect fractals, D is invariant to scale changes by

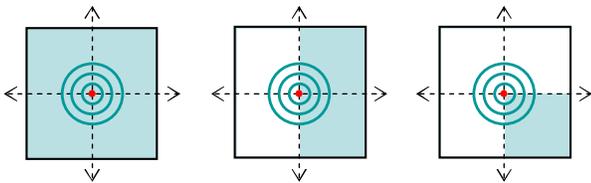


Figure 5. For uniform regions, edges and corners, $\mu(\bar{B}(\mathbf{x}, r))$ equals πr^2 , $\frac{\pi}{2} r^2$ and $\frac{\pi}{4} r^2$ respectively. The fractal dimension is 2 in each case. However, the fractal length is different and can therefore be used to distinguish between such texture primitives.

definition. In fact, D is invariant to bi-Lipschitz transformations of the image including affine and perspective distortions as well as certain non-linear mappings such as x^2 and e^x on the bounded interval [10, 100]. On the other hand, L has more discriminative power but in practise only rotational invariance. These points can be seen as follows (the proof proceeds along the lines of the one in [37]).

A bi-Lipschitz function f must be invertible and satisfy the constraint $c_1 \|x - y\| \leq \|f(x) - f(y)\| \leq c_2 \|x - y\|$ for constants $c_2 \geq c_1 > 0$ (essentially, smooth invertible mappings where both the original and the inverse function have bounded first derivative are bi-Lipschitz). Given a point \mathbf{x} in an image I for which the power law holds, i.e. $\log \mu_I(\bar{B}(\mathbf{x}, r)) = D \log r + L$, we would like to determine the invariance of L and D for a bi-Lipschitz transformed image $I'(f(\mathbf{x})) = I(\mathbf{x})$.

Let $\bar{B}(f(\mathbf{x}), r)$ be a closed disk or radius r centred around the transformed point $f(\mathbf{x})$ and $\log \mu_{I'} = D' \log r + L'$ be the log of its measure. Since f is invertible, there exists a point set P which is the pre-image of $\bar{B}(f(\mathbf{x}), r)$ and hence $\mu_I(P) = \mu_{I'}(\bar{B}(f(\mathbf{x}), r))$. The bi-Lipschitz constraint now ensures that $\bar{B}(\mathbf{x}, r/c_2) \subseteq P \subseteq \bar{B}(\mathbf{x}, r/c_1)$ so that $\mu_I(\bar{B}(\mathbf{x}, r/c_2)) \leq \mu_I(P) \leq \mu_I(\bar{B}(\mathbf{x}, r/c_1))$ since a subset’s measure must be less than or equal to the original set’s measure. Noting that $\mu_I(P) = \mu_{I'}(\bar{B}(f(\mathbf{x}), r))$ and substituting for the power law gives $D \log(\frac{r}{c_2}) + L \leq D' \log r + L' \leq D \log(\frac{r}{c_1}) + L$ which implies that $D' = D$ but $L' \neq L$. Thus the local fractal dimension is invariant to bi-Lipschitz transformations whereas the local fractal length is not.

Statistical characterisation In order to characterise textures statistically, the full PDF of the fractal features is modelled (details are given in Section 3). Figure 6 shows some texture images, and corresponding fractal dimension distributions. As can be seen, the distributions do not change much within a class and are yet distinct across all four classes (two different types of brick and two types of fabric). This is despite the fact that there is considerable intra-class variation including significant scale and perspective transformations as well as non-rigid surface deformations.

While distributions of local features are by now standard in texture classification [3, 8, 12, 18, 21, 27, 30, 31] we would still like to briefly point out their advantages over globally computed feature vectors. Firstly, feature distributions are more representative than global averages. For instance, most natural surfaces, and particularly inhomogeneous ones, will have an entire distribution of fractal dimensions corresponding to areas of different roughness rather than a global surface roughness parameter. Secondly, local distributions can be made relatively robust to real world effects such as shadowing or occlusion of small parts of the image. Thirdly, locally bi-Lipschitz invariant features can

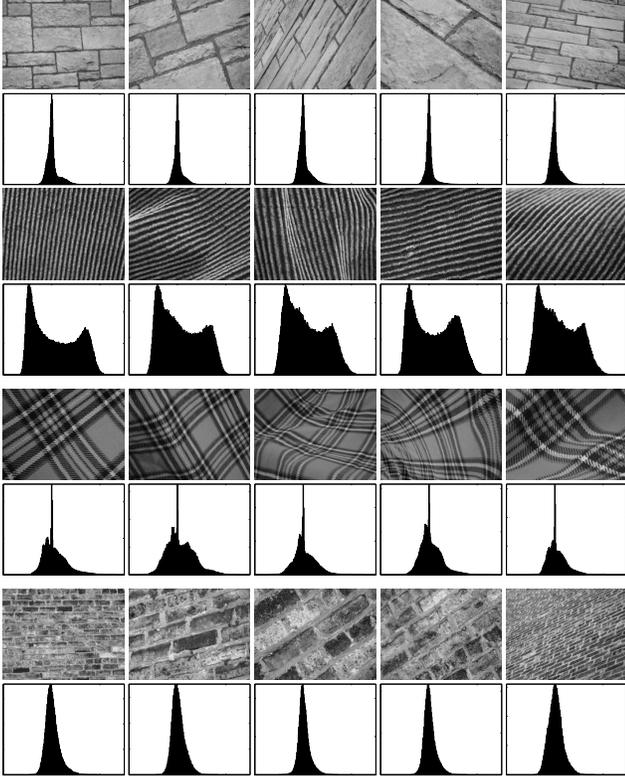


Figure 6. Images from four texture classes are shown along with their associated distributions of local fractal dimensions. Note that both classes of brick images have significant scale and perspective distortions while the two types of fabric also have considerable non-rigid surface deformation. Nevertheless, the densely sampled feature distributions are very similar within a class (due to bi-Lipschitz invariance) while also being easily distinguishable across the four classes.

account for more types of transformations than globally invariant ones. Finally, assuming that a fractal model holds locally is less restrictive than assuming it holds globally.

3. Texture Classification

In this section, we make concrete the local feature distributions and classifiers that are used. Experimental results are also presented on the UIUC and CURET texture databases.

Features Most fractal methods are based on measures calculated over filter responses. For instance, [37] generate features by defining measures over Gaussian, Laplacian and energy filters. Filters can smooth over image noise and lead to more robust features. However, they also have the drawback of lowering the level of bi-Lipschitz invariance. Note that this might not be a serious limitation in some cases as bi-Lipschitz invariant features are dependent on the rate of change of filter responses which might be less affected than

the absolute value of the filter responses themselves. Thus, basing fractal features on filter response measures depends on the specific application at hand. It should be avoided if there is a significant drop in the level of invariance. However, in our case, we also empirically verified that using filters does indeed lead to better classification performance.

We therefore characterise textures by the distribution of local fractal dimension and fractal length features obtained from multiple filter measures rather than just the single measure discussed so far. Defining $\mathbf{D}(\mathbf{x}) = [D_{\mu_1}(\mathbf{x}) \dots D_{\mu_n}(\mathbf{x})]$ and $\mathbf{L}(\mathbf{x}) = [L_{\mu_1}(\mathbf{x}) \dots L_{\mu_m}(\mathbf{x})]$ we estimate the joint density of \mathbf{D} , as well as that of \mathbf{L} . This boosts classification as observing how two measures covary provides more information than knowledge about how each varies independently.

Multiple measures are generated by first pre-filtering the images using the MR8 filter bank [31] and then applying the standard sum measure to each of the filter response images. The MR8 filter bank is a rotationally invariant, non-linear filter bank with 38 filters but only 8 filter responses. It contains edge and bar filters, each at 6 orientations and 3 scales, as well as a rotationally symmetric Gaussian and Laplacian of Gaussian filter (see Figure 7). Rotational invariance of the edge and bar filters is achieved by taking the maximum response over all orientations (more details can be found in [31]). Thus, in our case, $\mu_i(\bar{B}(\mathbf{x}, r)) = \sum_{\|\mathbf{y}-\mathbf{x}\| \leq r} |f_i(\mathbf{y})|$ where $f_i = \max_{\theta} F_i^{\theta} \star I$ for $1 \leq i \leq 8$ represents the i -th filter response image and the absolute value of filter responses has been taken to satisfy the measure requirement $\mu \geq 0$.

Each filter is made zero mean so as to be invariant to shifts in the illumination intensity. However, rather than achieve full affine illumination invariance by post-processing filter responses to have unit variance, we normalise by Weber's law [21, 31] instead as this is empirically found to give better results. Also based on empirical results, only five measures were used for calculating \mathbf{D} . The fractal dimensions calculated using the other three measures were found to be highly correlated and thus discarded. These measures correspond to the Gaussian, medium scale bar filter and smallest scale edge filter. However, all eight measures were used for calculating \mathbf{L} .

In summary, \mathbf{D} and \mathbf{L} are estimated at each pixel \mathbf{x} . The

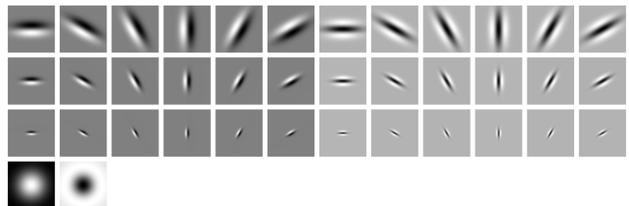


Figure 7. The MR8 filter bank.

local fractal dimension $\mathbf{D}(\mathbf{x})$ is a 5 dimensional vector invariant to geometric bi-Lipschitz transformations (modulo filter response variations) while the local fractal length $\mathbf{L}(\mathbf{x})$ is 8 dimensional and only rotation invariant. Both \mathbf{D} and \mathbf{L} are chosen to be invariant to local shifts in the illuminant intensity alone rather than full local affine illumination invariance.

Classification Our classification approach is standard and is based on nearest neighbour matching using a bag of visual words model. Owing to their different levels of invariance and discriminative power, fractal dimension and fractal length features are not combined into a single feature. Instead a separate classifier is learnt based on each feature individually – though the procedure for learning either is identical.

In the learning stage, either fractal dimension or fractal length features are computed densely from each training image. For every texture class in turn, features from a randomly chosen subset of training images are aggregated and clustered using the *K-Means* algorithm. The resultant cluster centres are known as textons and are aggregated over classes to form a dictionary of exemplar features. Given a texton dictionary, a model is learnt for a particular training image by labelling each of the image pixels with the texton that lies closest to it in feature space. The model is the normalised frequency histogram of pixel texton labellings. Each texture class is represented by a number of models corresponding to training images of that class which coarsely sample the imaging parameters.

In the classification stage, the set of learnt models is used to classify a novel image into one of the texture classes. This proceeds as follows: the fractal dimension or the fractal length features of the test image are generated and the pixels labelled with textons from the appropriate texton dictionary. Next, the normalised frequency histogram of texton labellings is computed. A nearest neighbour classifier is used to assign the texture class of the nearest model to the test image, where the distance between two normalised frequency histograms is measured using the χ^2 statistic, where $\chi^2(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \sum_i \frac{(x_i - y_i)^2}{x_i + y_i}$. In our experiments, we found that the size of the texton dictionary had very little affect on performance. The size of the dictionary was therefore set to about 2500 textons for both the UIUC and CURET databases.

Results on the UIUC database The UIUC database [20] contains 40 images each of 25 different texture classes thereby giving 1000 images in all (each image has resolution 640×480). The database represents a major improvement over the CURET textures [9] in that materials are imaged under significant viewpoint variations and some also have considerable surface deformations (see Figure 6

for examples). However, a drawback is that it is also much smaller than the CURET database, both in the number of classes as well as the number of images per class. It also has very few instances of a given material so that it is difficult to perform categorisation experiments [3, 12] or deduce generalisation properties of features. Furthermore, the high resolution of the images makes it unclear how features will performance in real world settings where textured regions on objects might be much smaller. Nevertheless, as far as scale and other viewpoint variations are concerned, the UIUC database is by far the most challenging and we therefore test the proposed features on it.

To assesses classification performance, M training images are randomly chosen per class while the remaining $40 - M$ images per class are taken to form the test set. Note that [20] present single descriptor results for the solitary case of $M = 10$ (all other results are for combinations of descriptors such as spin images and RIFT). We therefore implement their system so as to make comparisons as the training set size is varied. The best performance achieved by [20] for $M = 10$ is 90.15% (mean value over 200 splits) using spin image descriptors and combined classifiers based on Harris-affine and Laplacian blob detectors. Our implementation of their system achieves a comparable $90.17 \pm 1.11\%$.

M	\mathbf{D}	\mathbf{L}	[20]	[37]
20	95.40±0.92	94.96±0.91	93.62±0.97	93.04
15	94.09±0.98	93.66±0.96	92.42±0.99	91.11
10	91.64±1.18	91.25±1.13	90.17±1.11	88.79
05	85.35±1.69	84.96±1.66	84.77±1.54	82.99
	(a)	(b)	(c)	(d)

Table 1. UIUC results as the number of training images M is varied: (a) fractal dimension; (b) fractal length; (c) and (d) the methods of [20] and [37] respectively. Means and standard deviations have been computed over 1000 random splits of the training and test set.

Table 1 compares the performance of our local fractal dimension and length descriptors with the methods of [20] and [37]. As can be seen, the performance using either \mathbf{D} or \mathbf{L} is better than that achieved by the state-of-the-art fractal based method of [37]. Our results also compare favourably with that of [20]. For $M = 20$, local bi-Lipschitz invariant fractal dimension features achieve classification rates of $95.40 \pm 0.92\%$ while the affine adaption based spin image features of [20] achieve $93.62 \pm 0.97\%$. Local fractal length features achieve $94.96 \pm 0.91\%$. It is interesting to note that the performance of 5D fractal dimension features is better than that of the 100D spin images of [20]. This is due to the advantage of densely sampling features while avoiding characteristic scale selection. What is also curious is that 8D rotation invariant fractal length features also achieve such good performance despite the large scale

changes present in the UIUC dataset. However, this result is consistent with the trend in [28] where vanilla rotationally invariant patches were shown to give $97.83 \pm 0.63\%$ using standard nearest neighbour classifiers.

Results on the CURET database The CURET database [9] is a larger database and contains 61 texture classes with 205 images per class. However, the standard methodology on this database is to report results for 92 images per class. The remaining images do not have a sufficiently large portion of the texture visible to be cropped from the background and are therefore excluded.

There are a number of factors that make the CURET database challenging for a scale invariant method. To start with, the images in the database do not exhibit significant scale variation. As a result, scale invariant features tend to perform worse than features with just rotation or even no invariance but higher discriminative power. In addition, the images have low resolution (200×200 pixels) and are therefore not well suited to sparse interest point based methods. Nevertheless, one can expect to see such conditions in the real world and therefore a scale invariant method should still yield acceptable results on the CURET database if it is to be considered widely applicable.

M	D	L
46	96.12 ± 0.37	97.50 ± 0.30
23	92.50 ± 0.51	94.69 ± 0.45
12	86.70 ± 0.72	89.74 ± 0.66
06	78.05 ± 0.97	81.67 ± 0.96
	(a)	(b)

Table 2. CURET results as the number of training images M is varied: (a) fractal dimension and (b) fractal length. Means and standard deviations have been computed over 1000 random splits of the training and test set.

The evaluation methodology is similar to that used on the UIUC database. M images are randomly chosen per class for training, while the remaining $92 - M$ images per class are taken to form the test set. Table 2 presents the classification performance of the local fractal dimension and length features. For $M = 46$ training images, they achieve performances of $96.12 \pm 0.37\%$ and $97.50 \pm 0.30\%$ respectively. By contrast, the affine adaptation method of Lazebnik *et al.* using nearest neighbour classification achieves only $72.50 \pm 0.7\%$ [38]. This result is somewhat surprising but highlights the inadequacies of sparse, affine adaptation based methods. Even when multiple high dimensional descriptors are combined with multiple detectors and sophisticated SVMs employed, the affine adaptation results improve to only $95.30 \pm 0.4\%$ [38].

As is to be expected on the CURET database, rotation invariant fractal length feature achieve better results than the bi-Lipschitz invariant features. However, these are slightly

inferior to 97.64% which is the best rotation invariant result obtained by nearest neighbour classification on this database [29]. The best rotation invariant results for any classifier is 99.02% [2] when Gaussian Bayes classifiers are used.

4. Conclusions

In this paper, we have developed locally invariant, dense fractal features for the canonical texture classification problem. Features based on the fractal dimension are invariant to local bi-Lipschitz transformations and are robust as they avoid interest point detection and the affine adaptation process. Fractal length features are slightly more discriminative but possess only rotational invariance. Both features can be made invariant to affine illumination transformations if desired. On the UIUC database, classification based on these features yields results which are at least as good as leading fractal and affine adaptation based methods. However, on the CURET database, the performance of the proposed features is dramatically better. These results demonstrate that the fractal model is widely applicable to many texture classes and that fractal features provide a viable alternative to the affine adaptation process for dealing with scale and other viewpoint variations in textured images.

Yet, it should be noted that even though the best results are in the mid nineties or higher, the standard texture classification problem can not be considered solved or easy. In actuality, classification performance using only a few training image can drop quite dramatically even if there are only a small number of classes. Considerable progress therefore needs to be made in designing features and classifiers before performance can be considered acceptable. Furthermore, improved understanding of features and classifiers can only benefit both the categorisation problem as well as the simultaneous segmentation and classification problem in cluttered, real world scenes.

An area which we would like to explore as future work is to see if the fractal model can be made even more applicable. The sum measure used in this paper does not induce local fractal behaviour in all images. Even when it does, there is no guarantee that the resultant features are the most discriminative. It would therefore be preferable to actually learn a measure which induces fractal behaviour for the given classification task and leads to improved performance.

Acknowledgements

We are grateful to P. Anandan, Jitendra Malik and Andrew Zisserman for helpful discussions.

References

- [1] M. F. Barnsley and L. P. Hurd. *Fractal Image Compression*. 1992.

- [2] R. E. Broadhurst. Statistical estimation of histogram variation for texture classification. In *Proc. Intl. Workshop on Texture Analysis and Synthesis*, pages 25–30, Beijing, China, 2005.
- [3] B. Caputo, E. Hayman, and P. Mallikarjuna. Class-specific material categorisation. In *ICCV*, volume 2, pages 1597–1604, Beijing, China, 2005.
- [4] D. Charalampidis and T. Kasparis. Wavelet-based rotational invariant roughness features for texture classification and segmentation. *IEEE TIP*, 11(8):825–837, 2002.
- [5] B. B. Chaudhuri and N. Sarkar. Texture segmentation using fractal dimension. *IEEE PAMI*, 17(1):72–77, 1995.
- [6] D. Chetverikov and Z. Foldvari. Affine-invariant texture classification using regularity features. In M. Pietikainen, editor, *Texture Analysis in Machine Vision*. World Scientific, 2000.
- [7] F. S. Cohen and Z. Fan. Rotation and scale invariant texture classification. In *ICRA*, volume 3, pages 1394–1399, Philadelphia, PA, 1988.
- [8] O. G. Cula and K. J. Dana. 3D texture recognition using bidirectional feature histograms. *IJCV*, 59(1):33–60, 2004.
- [9] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real world surfaces. *ACM Transactions on Graphics*, 18(1):1–34, 1999.
- [10] S. Deguy, C. Debain, and A. Benassi. Classification of texture images using multi-scale statistical estimators of fractal parameters. In *BMVC.*, 2000.
- [11] K. J. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley, second edition, 2003.
- [12] E. Hayman, B. Caputo, M. Fritz, and J.-O. Eklundh. On the significance of real-world conditions for material classification. In *Proc. ECCV*, volume 4, pages 253–266, Prague, Czech Republic, 2004.
- [13] L. Kam and J. Blanc-Talon. Are multifractal multipermuted multinomial measures good enough for unsupervised image segmentation. In *CVPR*, volume 1, pages 58–63, Hilton Head Island, SC, 2000.
- [14] Y. Kang, K. Morooka, and H. Nagahashi. Scale invariant texture analysis using multi-scale local autocorrelation features. In *Int. Conf. on Scale Space*, pages 363–373, Hofgeismar, Germany, 2005.
- [15] L. M. Kaplan. Extended fractal analysis for texture classification and segmentation. *IEEE PAMI*, 8(11):1572–1585, 1999.
- [16] L. M. Kaplan and C. C. J. Kuo. Texture roughness analysis and synthesis via extended self similar (ESS) model. *IEEE PAMI*, 17(11):1043–1056, 1995.
- [17] J. M. Keller, S. Chen, and R. M. Crownover. Texture description and segmentation through fractal geometry. *CVGIP*, 45:150–166, 1989.
- [18] S. Konishi and A. L. Yuille. Statistical cues for domain specific image segmentation with performance analysis. In *CVPR*, volume 1, pages 125–132, Hilton Head, South Carolina, 2000.
- [19] P. Kube and A. Pentland. On the imaging of fractal surfaces. *IEEE PAMI*, 10(5):704–707, 1988.
- [20] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture representation using local affine regions. *IEEE PAMI*, 27(8):1265–1278, 2005.
- [21] T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textures. *IJCV*, 43(1):29–44, 2001.
- [22] G. McGunngile and M. J. Chantler. Evaluating Kube and Pentland’s fractal imaging model. *IEEE TIP*, 10(4):534–542, 2001.
- [23] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *IJCV*, 65(1/2):43–72, 2005.
- [24] S. Peleg, J. Naor, R. Hartley, and D. Avnir. Multiple resolution texture analysis and classification. *IEEE PAMI*, 6(4):518–523, 1984.
- [25] C. M. Pun and M. C. Lee. Log-polar wavelet energy signatures for rotation and scale invariant texture classification. *IEEE PAMI*, 25(5):590–603, 2003.
- [26] F. Schaffalitzky and A. Zisserman. Viewpoint invariant texture matching and wide baseline stereo. In *ICCV*, volume 2, pages 636–643, Vancouver, Canada, 2001.
- [27] C. Schmid. Weakly supervised learning of visual models and its application to content-based retrieval. *IJCV*, 56(1):7–16, 2004.
- [28] M. Varma and D. Ray. Learning the discriminative power-invariance trade-off. In *ICCV*, 2007.
- [29] M. Varma and A. Zisserman. A statistical approach to material classification using image patch exemplars. <http://research.microsoft.com/manik/pubs/varma07.pdf>.
- [30] M. Varma and A. Zisserman. Texture classification: Are filter banks necessary? In *CVPR*, volume 2, pages 691–698, Madison, Wisconsin, 2003.
- [31] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *IJCV*, 62(1–2):61–81, 2005.
- [32] J. L. Vehel. About lacunarity, some links between fractal and integral geometry, and an application to texture segmentation. In *ICCV*, pages 380–384, 1990.
- [33] J. L. Vehel. Fractal probability functions—an application to image analysis. In *CVPR*, pages 378–383, 1991.
- [34] J. L. Vehel, P. Mignot, and J. P. Berroir. Multifractals, texture, and image analysis. In *CVPR*, pages 661–664, 1992.
- [35] B. Wohlberg and G. de Jager. A review of the fractal image coding literature. *IEEE TIP*, 8(12):1716–1729, 1999.
- [36] Y. Xia, D. Feng, and R. Zhao. Morphology-based multifractal estimation for texture segmentation. *IEEE PAMI*, 15(3):614–623, 2006.
- [37] Y. Xu, H. Ji, and C. Fermuller. A projective invariant for textures. In *CVPR*, volume 2, pages 1932–1939, New York, New York, 2006.
- [38] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid. Local features and kernels for classification of texture and object categories: A comprehensive study. *IJCV*, 73(2):213–238, 2007.
- [39] J. Zhang and T. Tan. Affine invariant classification and retrieval of texture images. *PRL*, 36:657–664, 2003.