The Scale of Geometric Texture

Geoffrey Oxholm, Prabin Bariya, and Ko Nishino

Department of Computer Science
Drexel University, Philadelphia, PA 19104, USA
{gao25,pb335,kon}@drexel.edu

Abstract. The most defining characteristic of texture is its underlying geometry. Although the appearance of texture is as dynamic as its illumination and viewing conditions, its geometry remains constant. In this work, we study the fundamental characteristic properties of texture geometry—self similarity and scale variability—and exploit them to perform surface normal estimation, and geometric texture classification. Textures, whether they are regular or stochastic, exhibit some form of repetition in their underlying geometry. We use this property to derive a photometric stereo method uniquely tailored to utilize the redundancy in geometric texture. Using basic observations about the scale variability of texture geometry, we derive a compact, rotation invariant, scale-space representation of geometric texture. To evaluate this representation we introduce an extensive new texture database that contains multiple distances as well as in-plane and out-of plane rotations. The high accuracy of the classification results indicate the descriptive yet compact nature of our texture representation, and demonstrates the importance of geometric texture analysis, pointing the way towards improvements in appearance modeling and synthesis.

1 Introduction

With important applications in remote sensing, industrial surface inspection, and scene segmentation, texture analysis has been a popular area of study in computer vision for decades. Typically, this analysis is done on the 2D *appearance* of texture. Whereas the appearance of a texture depends greatly on illumination and viewing conditions, the underlying 3D *geometry* remains constant. We pursue a deeper understanding of texture geometry in order to enable important tasks, such as object recognition, to be done reliably in any condition. To do so, we study the two primary characteristics of geometric texture—self similarity and scale variability—and exploit them to perform surface normal estimation and geometric texture recognition.

To demonstrate the importance of texture geometry we first introduce a photometric stereo method specifically derived to leverage the repetitious nature of geometric texture. We observe that the appearance of each point is determined fundamentally by both the surface orientation at that point, and whether or not it is being directly illuminated. Since each of these factors cannot be estimated without knowledge of the other, we must estimate them jointly. To do so, we derive a probabilistic graphical formulation in which these two conditionally independent factors are linked by their combined effect on the appearance of each point. We sequentially order the dense set of observations as the light is waved over the texture. This allows us to derive novel priors for both factors.

A. Fitzgibbon et al. (Eds.): ECCV 2012, Part I, LNCS 7572, pp. 58–71, 2012.

In particular, we leverage the similarity of geometry that is present in spatially disjoint locations as a prior on the surface normals. Pixels whose intensities vary similarly as the light moves must have similar orientations. This increased contextual information allows us to minimize the effect of self-occlusion and interreflection.

Our second main contribution is a rotation-invariant scale-space representation of the geometry of texture. Our aim is to compactly encode how the apparent texture geometry varies with its distance from the observer so that a texture can be identified reliably regardless of its distance and orientation. Since the 3D surface geometry itself cannot be reliably recovered due to discontinuities of the surface, we directly use the surface normal field recovered with the proposed texture photometric stereo as the foundation for our representation. We first represent the surface normals using their spherical coordinates. This allows us to describe the geometry using a 2D histogram in which rotations of the texture correspond to shifts along the azimuthal axis. By applying the 2D Fourier transform we obtain a rotationally invariant representation of the texture geometry. Next, we compute the scale space of the texture geometry by convolving the surface normal field with Gaussian kernels of increasing standard deviations until the surface normal field is completely coherent. We show that this scale space of geometric texture accurately captures the evolution of the apparent surface normal field over increasing distance, and as such, approximates the underlying geometric scale space.

Our final contribution is a new texture database consisting of over $40,000$ images that we use to demonstrate the descriptiveness of our scale-space representation. We perform extensive classification experimentation with query textures at multiple distances, with varying degrees of in-plane and out-of-plane rotations as well as under simulated noise. In total, we test our representation on 600 combinations of distance and orientation. The high accuracy of the classification results indicate the descriptive yet compact nature of our texture representation, and demonstrates the importance of geometric texture analysis.

2 Related Work

Past work on texture analysis has focused primarily on the 2D imaged appearance of textures. In this domain, many methods have been proposed for rotation invariant and scale invariant classification [9,12]. Similarly, a number of methods have been proposed for scale-space analysis of texture appearance [7]. The appearance of texture, however, depends fundamentally on the underlying surface geometry. Unlike its appearance, the geometry of a texture does not change with the illumination and viewing conditions. In this work we focus directly on the geometry of surface texture and derive a rotationally invariant scale-space representation which we use to perform classification.

There has been some attention to illumination invariance. Smith [20] use photometric stereo to recover surface normals and then encode them as a surface orientation histogram. This approach, however, is geared at analysis to find surface defects, whereas our method is being used to perform classification. McGunnigle and Chantler [15] and Barsky and Petrou [2] extract and use explicit 3D shape and surface reflectance information to perform classification. Penirschke et al. [19] propose an illumination invariant classification scheme based on earlier work by Chantler et al. [6] wherein they study

the effect of changes in illumination direction on the image texture. Other approaches have involved explicitly training a classifier on multiple images of the texture from different viewpoints and illumination conditions [13, 22]. Dana and Nayar [8] model texture surface as a Gaussian-distributed random height field. To avoid this assumption we develop a novel scale-based representation in the frequency domain. Although some authors have considered the effect of the underlying geometry, by studying it directly we deepen the understanding of texture, and enable improvements to appearance modeling and synthesis.

In order to analyze the scale variability of geometric texture, dense and accurate surface normal information must be gathered. As first introduced by Woodham [23], the technique of photometric stereo enables the extraction of accurate surface orientations for simple Lambertian objects. As noted by Nayar et al. [16], problems arise when self-occlusions cause shadows, or interreflection exaggerates illumination conditions. Although there have been many advances in photometric stereo to these, and other ends, our method tackles the less-studied problem of estimating the geometry of textures.

To address self-occlusions, Barsky and Petrou [3] assume that each pixel is directly illuminated in at least two images. They then ignore light source directions where the observed intensity is low. Chandraker et al. [5] model light source visibility using a Markov Random Field with a strong spatial prior that encourages detection of large shadows. Similarly, Sunkavalli et al. [21], cluster pixels into "visibility subspaces" by observing the effect of large shadows. Nayar et al. [17] project a moving pattern onto the scene to separate the effect of direct illumination from that of interreflection and other "global" illumination phenomena. Wu and Tang [24] collect a dense set of images and use their inherent redundancy to detect and discard the observations that are most impacted by global illumination. We further exploit redundancy by leveraging the repetitious nature of texture. In a manner similar to that of Koppal and Narasimhan [11], we manually wave a light source over the texture and detect similarly oriented surface patches in physically separate locations based on the similarity of their appearance variation. Unlike Koppal and Narasimhan, we use the full appearance profile, not just extrema. This allows us to establish a non-parametric distribution of orientation vectors that we use as a prior in our expectation maximization framework.

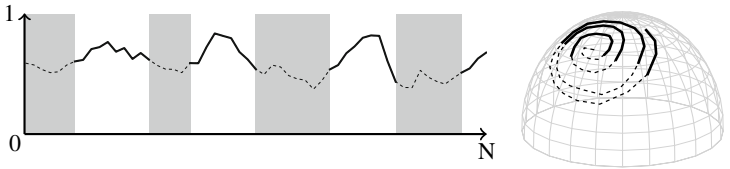
3 Texture Photometric Stereo

Textures, whether they are regular or stochastic, exhibit some form of repetition in their underlying geometry. We exploit this repetition in our texture photometric stereo in order to recover accurate surface normals.

3.1 Robust Estimation

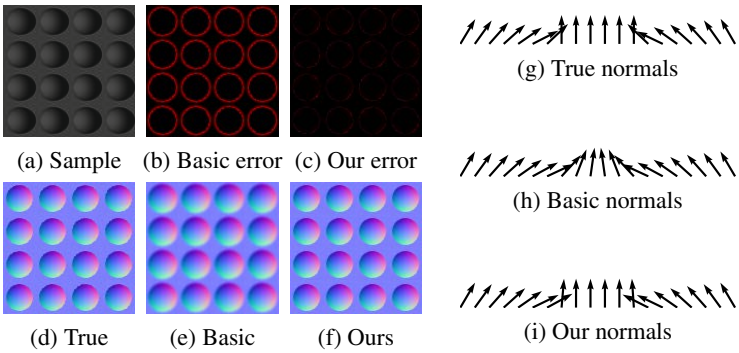
As light passes over a Lambertian texture the intensity I_t^x of a pixel x rises and falls depending on the relative angle of the light source direction $\boldsymbol{l}_t$ at that time t, and the surface orientation $\boldsymbol{n}^x$ at that point according to the basic Lambertian reflectance model

$$I_t^x = \rho^x \max(0, \boldsymbol{n}^x \cdot \boldsymbol{l}_t) \,, \tag{1}$$



(a) Observed intensities (b) Lighting directions

Fig. 1. Observed intensity profile and corresponding relative lighting direction at a single pixel. We use the intensity profile of a pixel (a) to estimate its light-source visibility. Shown as thick curves, convex regions of the (smoothed) intensity profile serve as good initial estimates of when a pixel is directly illuminated as they correspond to times when the light source passes near the surface normal (b).



(a) Sample (b) Basic error (c) Our error

(d) True (e) Basic (f) Ours

(g) True normals

(h) Basic normals

(i) Our normals

Fig. 2. Synthetic scene ground truth comparison. Without any model of light-source visibility, shadows smooth the recovered normal field (e). Using our method results in an accurate estimation (f). Brighter pixels in the error fields of (b) and (c) correspond to greater angular error. In the right half we show a slice of normals between two spheres. Note how the normals between the spheres in the basic method results (h) have been corrupted. Ray tracing interpolation is responsible for slight inconsistencies between our results and the ground truth.

where ρ^x is the albedo value at that point. Following Koppal and Narasimhan [11], as the light is waved over the texture, we take k discrete observations, and call $\boldsymbol{I}^x = \{I_t^x \mid t = 1, \ldots, k\}$ the *intensity profile* of pixel x. Fig. 1 shows the path of the moving light source on the unit hemisphere centered at a surface point. If we align the north pole of the hemisphere with the surface orientation at that point, then in accordance with Eq. 1, as the light moves closer to the pole, the observed intensity will increase.

Unfortunately, this model ignores a fundamental factor in appearance—light source visibility. Small scale self-occlusions lead to shadows in small regions and interreflection will increase the observed intensity. As shown in Fig. 2, failure to disregard these observations will result in inaccurate estimations of the underlying normal field. In our synthetic experimentation, our method results in a median error of $2.3°$, while the traditional method results in a median error of $14.8°$.

We use a binary visibility function V_t^x to indicate whether or not pixel x at time t is directly illuminated,

$$V_t^x = \begin{cases} 1 & \text{if pixel } x \text{ is directly illuminated} \\ 0 & \text{otherwise .} \end{cases} \tag{2}$$

By applying this function to both the observed intensity I_t^x, as well as the predicted intensity, we are able to effectively ignore misleading values. This augmented version of the Lambertian reflectance model may be written as

$$V_t^x I_t^x = V_t^x \rho^x \boldsymbol{n}^x \cdot \boldsymbol{l}_t \ , \tag{3}$$

where the $\max$ function from Eq. 1 has been absorbed into V. Since the surface normals are unit vectors, we simplify this notation by representing the albedo-scaled normal as $\boldsymbol{N}^x = \rho^x \boldsymbol{n}^x$, where $\rho^x = \|\boldsymbol{N}^x\|$.

A simple approach to estimating $\boldsymbol{V}$ would be to set $\boldsymbol{V}_t^x = 0$ for low intensity observations. For pixels with low albedo values, however, this is too limiting. A better approach is to use the portions of the light trajectory that are near the normal direction. As shown in Fig. 1a, we estimate these regions as convex portions of the intensity profile. To account for noise, we first smooth the intensity profile with a Gaussian distribution of standard deviation 3 before taking the second derivative with respect to t

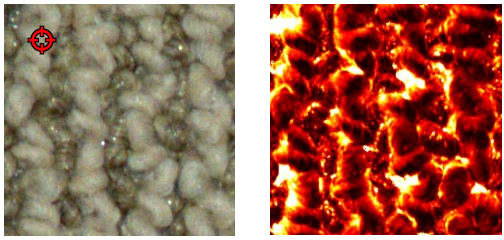
$$\boldsymbol{V}_t^x = \begin{cases} 1 & \text{if } (\mathcal{N}(3) * I^x)''(t) < 0 \\ 0 & \text{otherwise ,} \end{cases} \tag{4}$$

where $*$ denotes convolution. These regions, which are shown as thick solid lines in Fig. 1a, correspond well with lighting directions near the surface orientation (b).

3.2 Joint Visibility and Normal Estimation

Note that even under direct illumination, the effects of interreflection, subsurface scattering and other global illumination phenomena are impossible to avoid. Our goal is to utilize the observations where the appearance is dominated by the direct illumination component which will contain a more reliable encoding of the surface normal. With this in mind, estimating the visibility function is more akin to robust estimation wherein we seek to ignore as many outliers as possible while simultaneously avoiding conclusions based on sparse observations. Assuming known light source directions, estimating the optimal visibility function will then result in accurate surface normal estimations. In order to recover the visibility function, however, we must already know the correct surface orientations or we may be misled by low albedo values. In other words, since both the visibility function, and the surface normals contribute to the appearance of the texture, they must be estimated jointly.

To refine the visibility $\boldsymbol{V}$ and normal field $\boldsymbol{N}$ estimates we propose a probabilistic graphical model similar to a Factorial Markov Random Field [10]. We represent the visibility function and the surface normals as two separate, conditionally independent latent layers that are linked by their contribution to the observed appearance layer $\boldsymbol{I}$.



(a) Texture (b) Distance map

Fig. 3. Example distance map for one pixel. In (b) each pixel is given a brightness proportional to the distance of its intensity profile from that of the targeted pixel of (a) according to Eq. 8. Note how other downward-facing regions on the knots of carpet have lower (darker) distance values.

Using this formulation, we estimate the surface orientations and visibility functions jointly, by maximizing the posterior probability

$$p(\boldsymbol{N}, \boldsymbol{V} \mid \boldsymbol{I}) \propto p(\boldsymbol{I} \mid \boldsymbol{N}, \boldsymbol{V}) p(\boldsymbol{N}) p(\boldsymbol{V}) \,, \tag{5}$$

where $p(\boldsymbol{I}|\boldsymbol{N}, \boldsymbol{V})$ represents the likelihood of the observations given the estimated normals and visibility function, and the priors $p(\boldsymbol{N})$ and $p(\boldsymbol{V})$ encourage solutions that characterize the repetitious nature of the texture.

Likelihood. We assume that the noise inherent in the observations is normally distributed, with a common variance σ^2

$$p(\boldsymbol{I}|\boldsymbol{N}, \boldsymbol{V}) \propto \prod_t \prod_x \mathcal{N}\left(V_t^x I_t^x - V_t^x \boldsymbol{N}^x \cdot \boldsymbol{l}_t, \sigma^2\right) \,, \tag{6}$$

where $\boldsymbol{N}^x = \rho^x \boldsymbol{n}^x$. Note that without a prior on the surface orientation, pixels that are infrequently illuminated will be difficult to analyze. Additionally, a prior on the visibility function is necessary to avoid the trivial solution of $\boldsymbol{V} = \boldsymbol{0}$.

Surface Normal Prior. We exploit the repetitious nature of texture with a novel non-parametric prior across geometrically similar, but spatially disjoint locations. As noted above, if two pixels share similar intensity profiles, they must share similar surface orientations [11]. We translate this observation into a prior on the surface normals that encourages similar orientation for pixels with similar intensity profiles. We let Ω^x be the set of pixels in the spatially disjoint *intensity profile cluster* of x, and formulate the prior using a kernel density as

$$p(\boldsymbol{N}) = \prod_x \frac{1}{|\Omega^x|} \sum_{y \in \Omega^x} \mathcal{N}\left(\angle(\boldsymbol{N}^x, \boldsymbol{N}^y), h\right) \,, \tag{7}$$

where $|\Omega^x|$ is the set's cardinality, $\mathcal{N}$ is the normal density function, with variance h, and $\angle(\cdot, \cdot)$ is the angle between the two normals.

We let Ω^x be the set of locations $\Omega^x = \{y\}$ for which $\mathrm{dist}(x, y)$ is less than some threshold. We determine the intensity profile distance between two locations using only

the portions of their intensity profiles for which both locations are directly illuminated, i.e. $\{t \mid V_t^x = V_t^y = 1\}$. More precisely, we compute the distance using the cosine similarity of the visibility-clamped intensity profiles

$$\mathrm{dist}(x,y) = 1 - \frac{(\boldsymbol{V}^x \circ \boldsymbol{V}^y \circ \widetilde{\boldsymbol{I}}^x) \cdot (\boldsymbol{V}^x \circ \boldsymbol{V}^y \circ \widetilde{\boldsymbol{I}}^y)}{\|\boldsymbol{V}^x \circ \boldsymbol{V}^y \circ \widetilde{\boldsymbol{I}}^x\| \, \|\boldsymbol{V}^x \circ \boldsymbol{V}^y \circ \widetilde{\boldsymbol{I}}^y\|}\,, \tag{8}$$

where $\circ$ denotes element-wise multiplication, i.e. visibility-clamping, $\widetilde{\boldsymbol{I}}_t^x = \boldsymbol{I}_t^x / \rho_x$, and $\|\cdot\|$ indicates the $L2$ norm. By dividing by the estimated albedo, we allow pixels with different albedos, but similar orientations to be part of the same cluster.

Fig. 3 shows the distance map for one pixel that is located on the downward-facing slope of a carpet knot. As can be seen by the dark regions, this pixel has a low distance value to similarly oriented surface patches. The bright regions, on the other hand, correspond to areas that are not frequently illuminated at the same time, or seldom have similar intensity values.

Visibility Function Prior. As the light source is waved across the texture, each pixel will remain illuminated for consecutive observations, and in some consistency with the pixels surrounding it. In other words, the visibility function $\boldsymbol{V}$ should be piecewise constant both spatially, as well as temporally. Further, we impose a unary prior on the entire visibility function to avoid the trivial solution of $\boldsymbol{V} = \boldsymbol{0}$. The combined prior is

$$\begin{aligned} p(\boldsymbol{V}) = &\exp\big(-\beta_u^V \, \|\boldsymbol{1} - \boldsymbol{V}\|_1\big) \\ &\exp\Big(-\beta_s^V \sum_t \sum_x \sum_{y \in \mathsf{N}_s^x} |V_t^x - V_t^y|\Big) \\ &\exp\Big(-\beta_t^V \sum_t \sum_x \sum_{y \in \mathsf{N}_t^x} |V_t^x - V_t^y|\Big)\,, \end{aligned} \tag{9}$$

where N_s^x indicates the four spatial neighbors of pixel x, N_t^x indicates the two temporal neighbors (i.e. the previous and subsequent observations) and $\beta_{\{u,s,t\}}^V$ are the relative weights of the unary, spatial, and temporal components.

Note that setting the spatial prior weight β_s^V quite high, and the temporal prior β_t^V to zero, would result in a similar formulation to that of Chandraker et al. [5] who assume self-occlusions cast large shadows. In our case, however, we use a weak spatial prior to allow for small-scale detail.

Expectation Maximization. We estimate the visibility function and surface normals using expectation-maximization. In the expectation step we iterate back and forth, updating each layer while considering the other as a *pseudo-observable* [10]. In both cases, we can find the maximum a posterior by minimizing its negative log likelihood. In the case of the visibility function, since each node can take on only two possible states (0 or 1), we can approximate the minimum quickly using graph cuts [4].

Due to the nonparametric nature of our surface normal prior, we must make an approximation to avoid intractability. Specifically, we use the previously estimated values for the surface normals in the cluster $\boldsymbol{N}^{y \in \Omega^x}$ in Eq. 7. This allows each surface normal

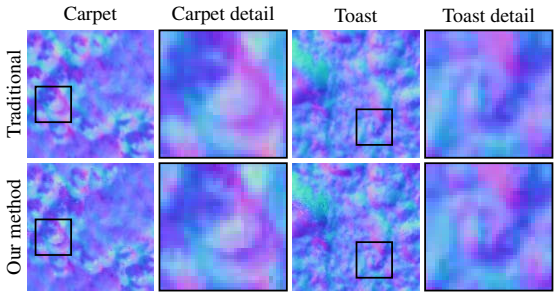


Fig. 4. Resulting normal field comparison. Note the increased detail of our method.

to be updated independently of the others. To minimize the negative log posterior for each pixel we use a gradient descent method.

In the maximization step, we update the noise variance σ^2 using the maximum likelihood estimator. These two steps, updating the latent values, and maximizing the likelihood, are iterated until convergence which we define as an average change in surface orientation of less than 1°.

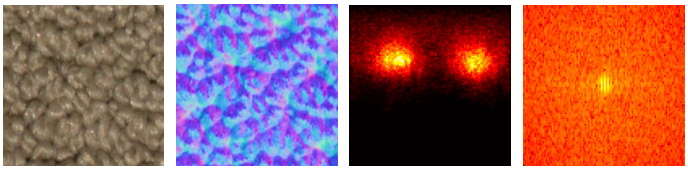
In Fig. 2 we show the result of our method on a synthetic scene of several spheres embedded in a plane. In addition to self-occlusion, as the light passes over the scene, each sphere casts large shadows on the background plane. These shadows result in erroneous normal estimations around each sphere when the traditional approach is used. On the right hand side of the figure we show a series of normal vectors from one slice of the resulting normal fields. Note how our method is able to accurately recover the sharp change from the background plane to the sphere. In Fig. 4 we compare our result for two real-world textures to the result using the traditional method. Qualitatively, one can see the increased clarity that results from our method over the traditional approach.

4 Encoding Geometric Texture

With the exception of sharp depth discontinuities, the geometry of texture is well described by the normal field. This normal field, however, will change dramatically as the texture is rotated, or moved towards or away from the viewer. An ideal representation of texture geometry will be both rotation-invariant and will encode its scale variability.

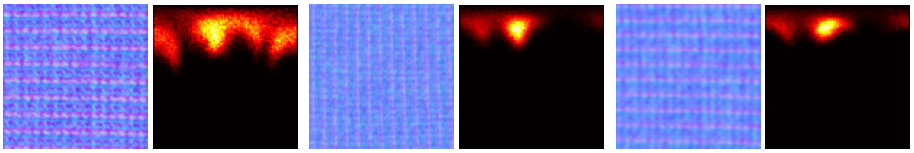
4.1 Rotation-Invariant Base Representation

In order to obtain a rotation-invariant representation of texture geometry, we first adjust the normal field so that the global surface normal (i.e., the mean normal) is aligned with the positive z-axis. This accounts for small out-of-plane rotations of the texture. As shown in Fig. 5c, we then form a normalized two-dimensional histogram $h(\phi, \theta)$ from the azimuth angle ϕ and polar angle θ of each surface normal. For our experiments, we quantize the polar and azimuth angles into 100 bins. In Fig. 5c we show one such histogram in which brighter regions correspond to higher values. We can see that the texture is relatively coarse by observing that it has a highly structured grouping



(a) Carpet texture (b) Normal field (c) Histogram (d) FFT amp.

Fig. 5. Steps to form base representation. After computing the surface normals (b), we build a 2D histogram (c) from the polar and azimuth angles of each normal vector. We use the rotation-invariant amplitude of the frequency domain (d) as our base representation.



(a) Near (b) Far (c) Estimated

Fig. 6. Comparison of real and estimated scale. When a texture (in this case, a cloth) is imaged at an increased distance, the apparent normal field becomes more coherent (b). We approximate this by smoothing the original normal field (a). The resulting histogram () is quite similar to the histogram of the observed texture.

of normals. In this histogram, a rotation about the mean normal corresponds to a shift along the horizontal (azimuth) axis. The final step in our base representation is to apply Fourier transform to the histogram. The amplitude of the transform, shown in log-space in Fig. 5d, is then a rotation invariant characterization of the normal field.

We define the similarity of two such frequency representations using the Jensen-Shannon divergence [14]. Given two surface textures A and B, we first compute their respective rotation-invariant representations $H_A(u, v)$ and $H_B(u, v)$. We then define their similarity as the Jensen-Shannon divergence of H_A and H_B

$$\mathcal{J}(H_A, H_B) = \frac{1}{2}\left(\mathcal{K}(\widetilde{H}_A, C) + \mathcal{K}(\widetilde{H}_B, C)\right) , \tag{10}$$

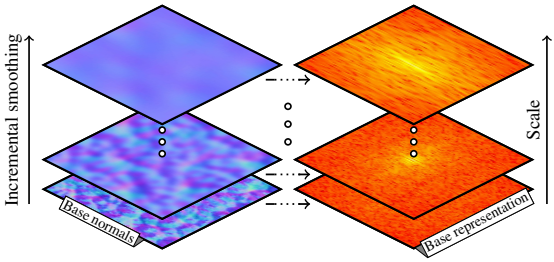
where $\widetilde{H}_A$ and $\widetilde{H}_B$ are the normalized distributions

$$\widetilde{H}_A = \frac{H_A}{\sum\sum |H_A|} \quad \widetilde{H}_B = \frac{H_B}{\sum\sum |H_B|} , \tag{11}$$

C is their mean, and $\mathcal{K}$ is the Kullback-Leibler divergence.

4.2 Scale-Space Representation

In Fig. 6 we show the surface normal field for a cloth at two distances. On the far left, the cloth is relatively close to the camera, and in the middle the cloth has been moved twice



(a) Normal fields (b) Our representation

Fig. 7. Building of scale-space representation. We represent the scale space of geometric texture with frequency histograms (b) built from the polar histogram (not shown) of the incrementally smoothed normal field (a).

the distance away. Although the geometry of the cloth has not changed, the number of microfacet surface normals subtended by each pixel has increased. In other words, as scale increases, the appearance of each pixel is determined by an increasing number of surface orientations. In effect, the average surface normal captured by a given pixel is computed across a greater area as scale increases. This observation, also underlying the recent work on geometric scale space of range images [1, 18], motivates our approach to modeling the scale variability of geometric texture.

We describe the scale space of geometric texture by filtering the surface normal field with Gaussian kernels of increasing standard deviation. In Fig. 6 we illustrate the result of this approach by comparing the resulting normal histogram (c) with that of the texture at a greater distance (b). The result is quite similar to the observed histogram.

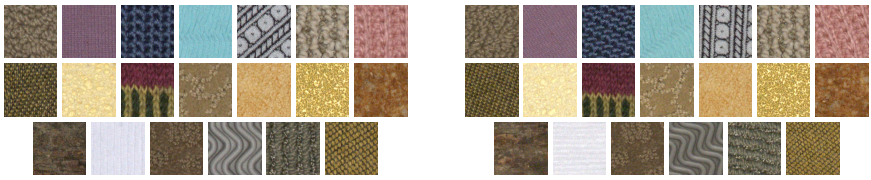
Fig. 7 summarizes the creation process of our scale-space representation. Starting in the lower left, we begin with the observed normal field of texture. As shown in (a), we then smooth this normal field at small intervals. At each interval, the polar histogram is formed and its Fourier transform is computed (b). This process is repeated until the normal field is sufficiently coherent (has an angular standard deviation of 2°). Note that as this happens, the polar histogram becomes nondescript, and corresponding frequency image flattens out, as can be seen at the top of the stack. Smooth textures, whose normal fields are already quite coherent, require few scales to describe their scale variability, while coarse textures will have larger stacks.

5 Classification

We evaluate the effectiveness of our rotation-invariant scale-space representation of texture geometry by performing classification on a new database.

5.1 Classifying a Query Texture

When a query texture is imaged, and its normal field is estimated, classification is performed on its frequency amplitude histogram. In order to find the correct texture, we must also find the correct scale. One approach would be to compute the similarity (Eq. 10) between the query texture and every scale of every database texture. Needless



(a) Base distance, no rotation (b) Assorted distances and rotations

Fig. 8. Select database samples. Each of our 20 textures is imaged with 30 combinations of orientation and distance.

to say, this process quickly becomes too computationally expensive as the number of textures and scales increases. To more quickly estimate the likely textures we first compare the total energy of the query texture to that of each database texture at each scale. The total energy is computed as the sum of the unnormalized frequency amplitudes.

5.2 Geometric Texture Database

As depicted in Fig. 8, we introduce a new database that covers 20 textures at different distances, with different in-plane and out-of-plane rotations[1]. To our knowledge, this is the only public database that offers multiple distances for each texture in addition to multiple in-plane and out-of-plane rotations. Our database, which contains stochastic as well as regular textures, includes organic materials such as bark and toast, woven materials such as cloth crocheted fabric, and synthetic materials such as a carpet and sandpaper. For practical reasons, we did not construct a lighting apparatus for each distance. Instead, each dataset was lit by manually waving a light source around in a spiral pattern like the one shown in Fig. 1b. The lighting direction was estimated using a reflective sphere. Each dataset consists of approximately 65 images. Every effort was made to keep the distance from the subject constant, but some amount of light attenuation will necessarily be present. To address this, each image is white-balanced using the light source direction and a white paper in the scene. Since each texture is imaged with more than $2,000$ different combinations of orientation, distance, and lighting direction, it is also useful for appearance modeling.

6 Experimental Validation

Distance. For each texture in the model database, acquired at a distance D1, we acquired additional real data to be used as query instances for classification at two increasing distances D2 and D3. All 20 query textures at D2 were classified correctly, while 19 were correctly classified at D3. These are the the starting points in the top left of Fig. 9.

Noise. Fig. 9 shows the robustness of our method to increasing levels of Gaussian noise in the normal field. The horizontal axis corresponds to the standard deviation of the Gaussian noise applied to each dimension of normal field. Below the graph are sample

[1] This database is available on-line at `http://cs.drexel.edu/~kon/texture`

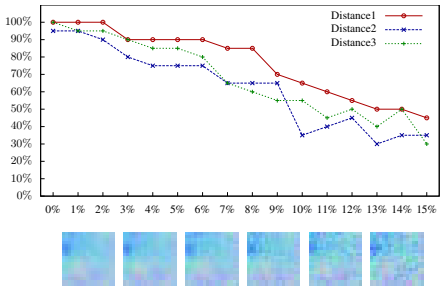
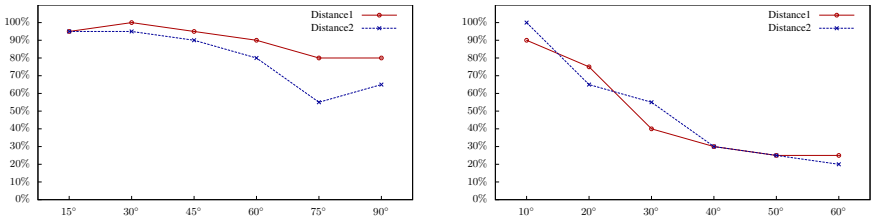


Fig. 9. Recognition rates under increasing Gaussian noise. The horizontal axis is the standard deviation of Gaussian noise added to the normal field. A sample region is shown for every third level of noise to illustrate the degree of corruption. Note that high accuracy is achieved even with significant corruption.



(a) In-plane rotation

(b) Out-of-plane rotation.

Fig. 10. Recognition rates for rotations Out-of-plane rotations expose previously unseen structures, causing novel normal fields which are more challenging to classify. Although recognition rates decline slightly at increased angles, high accuracy is persistent.

patches of a normal field at every third level of noise corruption. The fourth sample patch shows the high degree of noise at $\sigma = 9\%$. As shown in red, at this level 70% of the original textures are classified correctly. The blue and green lines correspond to the increased distances of D2 and D3, respectively. Since each surface normal captures an increased amount of surface area at these distances, noise has a greater effect.

In-Plane Rotation. At distances D1 and D2 we rotated each texture 6 times in increments of $15°$, stopping at a rotation of $90°$. Moving beyond $90°$ would be redundant, since the same result can be achieved by rotating the images $90°$. Overall, at distance D1 we correctly classify 88% of the 120 query textures, and at distance D2 we correctly classify 81% of the query textures. In Fig. 10a we show the results for each rotation increment.

Out-of-Plane Rotation. At distances D1 and D2 we performed out-of-plane rotations as well. These were done in increments of $10°$, stopping at $60°$. The results are summarized in Fig. 10b. Out-of-plane rotations present a unique challenge in that the rotations expose geometry that was previously hidden. In essence, out-of-plane rotations present novel textures as the side of the original texture becomes visible.

Arbitrary Rotations. At distance D3 we tested 3 arbitrary combinations of in-plane and out-of-plane rotations for each texture. Table 1 shows the results of all 60 tests.

Table 1. Recognition rates for combinations of in-plane and out-of-plane rotations. For quick inspection, higher recognition rates are given a brighter background coloring. The last row and column are sums. Note the high success rates for the top row and upper left region.

Out-of-plane rotation angle \ In-plane rotation angle	15°	30°	45°	60°	75°	90°	Σ
10°	3/3	5/5	3/5	3/4	1/1	1/2	16/20
20°	4/5	2/2	1/2	1/2	2/6	1/3	11/20
30°	0/4	3/4	1/2	1/2	2/4	4/4	11/20
Σ	7/12	10/11	5/9	5/8	5/11	6/9	38/60

Each row corresponds to an increased out-of-plane rotation angle, while each column corresponds to an increased in-plane rotation angle. Each entry in the table shows the total correct classifications and the number of experiments run with that combination of in-plane and out-of-plane rotation angles as a fraction. The shading of each cell indicates the overall success rate. Note that successful cases tend to have lower out-of-plane rotations, as expected.

7 Conclusion

In this work we have directly studied the geometry of texture. Through careful analysis of its two key characteristics—self similarity and scale variability—we have derived a photometric stereo method specifically tailored to exploit the repetitive nature of texture geometry, and have introduced a compact representation for the scale space of geometric texture. To evaluate our methods we have also introduced an extensive new texture database. Experimentally, we have shown that our compact scale-space representation is highly discriminative. We believe these results provide a sound foundation to explore the use of geometric texture in longstanding problems in computer vision, including exploiting texture geometry to aid object recognition and scene understanding.

Acknowledgments. This work was supported in part by the Office of Naval Research grant N00014-11-1-0099, and the National Science Foundation awards IIS-0746717 and IIS-0964420.

References

1. Bariya, P., Novatnack, J., Schwartz, G., Nishino, K.: 3D Geometric Scale Variability in Range Images: Features and Descriptors. Int'l Journal of Computer Vision 99, 232–255 (2012)
2. Barsky, S., Petrou, M.: Classification of 3D Rough Surfaces Using Color and Gradient Information Recovered by Color Photometric Stereo. In: SPIE Conf. on Visualization and Optimization Techniques (2001)
3. Barsky, S., Petrou, M.: The 4-Source Photometric Stereo Technique For Three-Dimensional Surfaces in the Presence of Highlights and Shadows. IEEE Trans. on Pattern Analysis and Machine Intelligence 25(10), 1239–1252 (2003)

4. Boykov, Y., Kolmogorov, V.: An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision. IEEE Trans. on Pattern Analysis and Machine Intelligence 26(9), 1124–1137 (2004)
5. Chandraker, M., Agarwal, S.: ShadowCuts: Photometric Stereo with Shadows. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition, pp. 1–8 (2007)
6. Chantler, M., Schmidt, M., Petrou, M., McGunnigle, G.: The Effect of Illuminant Rotation on Texture Filters: Lissajous's Ellipses. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part III. LNCS, vol. 2352, pp. 289–303. Springer, Heidelberg (2002)
7. Cohen, F., Fan, Z., Patel, M.: Classification of Rotated and Scaled Textured Images Using Gaussian Markov Random Field Models. IEEE Trans. on Pattern Analysis and Machine Intelligence 13(2), 192–202 (1991)
8. Dana, K.J., Nayar, S.K.: 3D Textured Surface Modeling. In: IEEE Workshop on the Integration of Appearance and Geometric Methods in Object Recognition, pp. 44–56 (1999)
9. Fountain, S.R., Tan, T.N., Baker, K.D.: A Comparative Study of Rotation Invariant Classification and Retrieval of Texture Images. In: British Machine Vision Conference, pp. 266–275 (1998)
10. Kim, J., Zabih, R.: Factorial Markov Random Fields. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part III. LNCS, vol. 2352, pp. 321–334. Springer, Heidelberg (2002)
11. Koppal, S., Narasimhan, S.: Clustering appearance for scene analysis. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 1323–1330 (2006)
12. Lazebnik, S., Schmid, C., Ponce, J.: Affine-Invariant Local Descriptors and Neighborhood Statistics for Texture Recognition. In: IEEE Int'l Conf. on Computer Vision (2003)
13. Leung, T., Malik, J.: Recognizing Surfaces Using Three-Dimensional Textons. In: IEEE Int'l Conf. on Computer Vision, pp. 1010–1017 (1999)
14. Lin, J.: Divergence Measures Based on the Shannon Entropy. IEEE Trans. on Information Theory 37(1), 145–151 (1991)
15. McGunnigle, G., Chantler, M.: Rough Surface Classification Using Point Statistics from Photometric Stereo. Pattern Recognition Letters 21, 593–604 (2000)
16. Nayar, S., Ikeuchi, K., Kanade, T.: Shape from Interreflections. Int'l Journal of Computer Vision 6(3), 173–195 (1991)
17. Nayar, S., Krishnan, G., Grossberg, M.D., Raskar, R.: Fast Separation of Direct and Global Components of a Scene using High Frequency Illumination. ACM Trans. on Graphics 25(3), 935–944 (2006)
18. Novatnack, J., Nishino, K.: Scale-Dependent 3D Geometric Features. In: IEEE Int'l Conf. on Computer Vision, pp. 1–8 (2007)
19. Penirschke, A., Chantler, M., Petrou, M.: Illuminant Rotation Invariant Classification of 3D Surface Textures Using Lissajous's Ellipses. In: Intl. Workshop on Texture Analysis and Synthesis (2002)
20. Smith, M.: The analysis of surface texture using photometric stereo acquisition and gradient space domain mapping. Image and Vision Computing 17(14), 1009–1019 (1999)
21. Sunkavalli, K., Zickler, T., Pfister, H.: Visibility Subspaces: Uncalibrated Photometric Stereo with Shadows. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part II. LNCS, vol. 6312, pp. 251–264. Springer, Heidelberg (2010)
22. Varma, M., Zisserman, A.: Classifying Images of Materials: Achieving Viewpoint and Illumination Independence. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part III. LNCS, vol. 2352, pp. 255–271. Springer, Heidelberg (2002)
23. Woodham, R.J.: Photometric Method for Determining Surface Orientation from Multiple Images. Optical Engineering 19(1), 139–144 (1980)
24. Wu, T.-P., Tang, C.-K.: Photometric Stereo Via Expectation Maximization. IEEE Trans. on Pattern Analysis and Machine Intelligence 32(3), 546–560 (2010)