In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) 2001*

Compact Representation of Bidirectional Texture Functions

Oana G. Cula
CS Department

Kristin J. Dana
ECE Department

Rutgers University
Piscataway, NJ 08854
{*oanacula,kdana*}*@caip.rutgers.edu*

Abstract

A bidirectional texture function (BTF) describes image texture as it varies with viewing and illumination direction. Many real world surfaces such as skin, fur, gravel, etc. exhibit fine-scale geometric surface detail. Accordingly, variations in appearance with viewing and illumination direction may be quite complex due to local foreshortening, masking and shadowing. Representations of surface texture that support robust recognition must account for these effects. We construct a representation which captures the underlying statistical distribution of features in the image texture as well as the variations in this distribution with viewing and illumination direction. The representation combines clustering to learn characteristic image features and principle components analysis to reduce the space of feature histograms. This representation is based on a core image set as determined by a quantitative evaluation of importance of individual images in the overall representation. The result is a compact representation and a recognition method where a single novel image of unknown viewing and illumination direction can be classified efficiently. The CUReT (Columbia-Utrecht reflectance and texture) database is used as a test set for evaluation of these methods.

1 Introduction

Recent trends are redefining the term texture in computer vision and graphics research. In computer vision, texture recognition algorithms were typically tested on simple databases like Brodatz [1] which do not consider variations of texture with viewing and illumination direction. In computer graphics, textured objects were synthesized by warping a single texture image onto a 3D object. Traditionally, texture referred to image texture, i.e. random or regular variations of pixel intensities. Figure 1 illustrates the large variation in image texture as a function of viewing and illumination direction. The sample shown exhibits surface texture, which can be defined as random or regular variations of surface geometry and/or reflectance. Indeed, for recognition tasks, surface texture is inherently more interesting since it reveals information about the imaged scene invariant of the imaging parameters. Most natural surfaces exhibit fine-scale geometry that cannot easily be measured and yet contributes significantly to overall appearance. Useful computational representations are constructed from measurable image texture, but the goal is to characterize the underlying surface texture.

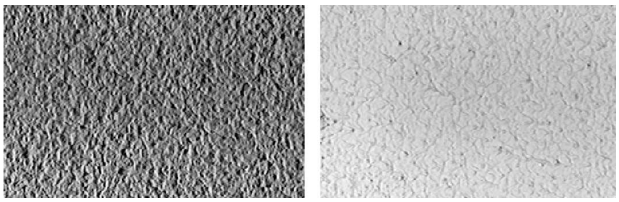

Figure 1: Two instances of the same plaster sample, captured under different imaging conditions. The appearance changes significantly with changes in viewing and illumination direction.

In [5][6], the term bidirectional texture function (BTF) is introduced to describe the variation of image texture with viewing and illumination direction. The more familiar term bidirectional reflectance distribution function (BRDF) is used to describe reflectance as a function of these imaging parameters. The BTF is simply a BRDF that also varies spatially as a function of the Euclidean coordinates on a planar surface patch. Because the BTF is image texture, it is an observable or measurable quantity. Of course, the specific pixel intensity values in the observed image are not significant for building characteristic surface representations since these values change with different instances of the same type of texture. Instead, statistical properties can be extracted from BTF's in order to construct a useful representation for the entire surface class.

BTF-based modeling is an integral part of developing methods for recognition tasks. In this work we develop a compact and practical representation that supports recognition of a surface using a single image from an unknown viewing and illumination direction. Only a few prior methods for BTF-based texture recognition have been introduced in the literature. Many of these methods use the CUReT database [6] as a source of test images.

The BTF-based texture recognition method in [8][9] defines and uses 3D textons. In this method, the images in the BTF database are aligned spatially so that the output of a filter bank applied to the BTF image set forms an ap-

pearance vector. This appearance vector encodes the filter responses at a pixel as a function of viewing and illumination direction. These aggregate vectors of filter outputs are clustered to identify key groups termed 3D textons. The histograms of 3D textons is the texture class representation for the surface. Another BTF-based recognition method [4] computes multiscale gradient histograms for each image in the BTF. Principal components analysis (PCA) is performed on the histograms as a function of viewing and illumination direction as in the methods of appearance-based object recognition [12][10][11]. Our recognition method has some similarities with the method in [9] and the method in [4]. Specifically, our method uses clustering on filter outputs to identify image features. However in our approach, clustering is done on the image filter outputs not the aggregation of filter outputs over multiple viewing/illumination directions. This approach leads to two key practical advantages. First, the images of the BTF need not be aligned. Second, the illumination and viewing direction of the images need not be known. For our approach, the histograms of image textons as a function of imaging parameters is the surface representation and PCA is used to reduce the dimensionality of the histogram space. Other BTF representations in the literature include [13] which uses a color correlation model to classify images of textured surfaces. The use of color is a strong cue for the CUReT sample set, because many of the samples are well separated in color space. In this work, we are interested in recognition based on gray scale image features so that two samples with different surface texture can be distinguished even if their colors are very similar.

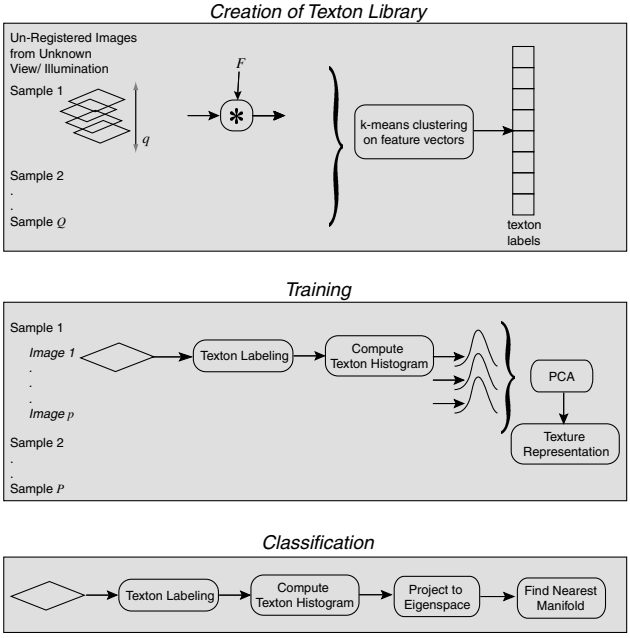

Figure 2: Three steps of BTF-based recognition. First, an image texton library is obtained by grouping image features from a set of samples used for learning. Second, PCA is done on the histograms of image textons using the SLAM software library [12]. Finally, recognition is accomplished by finding the closest sample to the novel point in the texton histogram eigenspace.

2 BTF-based Recognition

Modeling surface detail based on appearance of textured surfaces under varied imaging conditions is a convenient alternative to modeling fine scale geometry and reflectance of these surfaces. Therefore in our work we design a texture recognition technique based on the BTF, which captures the complex appearance of the surface under different viewing and illumination directions. Our BTF-based recognition method has three steps as shown in Figure 2. First, an image texton library is obtained by grouping image features from a set of samples used for learning. Second, PCA is done on the histograms of image textons to obtain a representation of the texture class. Finally, classification is accomplished by finding the closest sample to the novel point in the texton histogram eigenspace.

2.1 Image Texton Library

The utility of a characteristic texton library is emphasized in [9]. To build our version of a texton library, texture images from Q samples corresponding to q specific imaging conditions are analyzed by filtering to capture the apparent local image structure as illustrated in Figure 2.

We use a multiscale filter bank F, consisting of oriented Gaussian derivative filters, center surround derivative filters, and low-pass Gaussian filters, on three different scales. So at each pixel in each texture image there is a feature vector composed of the filter responses. Due to local repetitiveness in the texture image, the feature vectors give an overly redundant description of the local structure in the texture image. Therefore, we perform feature grouping, and the result is a set of structural appearance prototypes, termed image textons, where the term texton is introduced as the element of texture perception in [7]. The global set of feature groups is formed using the filter outputs from a collection of textured surface samples and multiple images per sample. Because of similarities across different textures, we can employ the image texton library for the purpose of recognizing novel textured surfaces. That is, the set of textures involved in creating the texton library, and the set of textures used for recognition, are disjoint. For feature grouping, we employ the k-means algorithm, a greedy, iterative and supervised clustering procedure, which makes use of the first order statistics of the data, and finds a predetermined number of centers in the data space, by minimizing the sum-of-squared errors criterion. Our filtering procedure and clustering is the same as [9], with the exception that the vectors for feature grouping are filter outputs from individual images not the aggregate of filter outputs as a function of viewing and illumination direction. Therefore while [9] uses the term 3D texton, we use the term image texton.

The motivation for using image textons is described in detail in [2] and is summarized here. For the 3D texton method of [9], variations with imaging conditions are accounted for in the local feature. But, in the image texton method the local feature is not a function of the imaging parameters. Instead the distribution of image textons varies with imaging parameters. This difference has consequences in terms of the amount of information needed

Figure 3: The collection of 20 texture samples employed in constructing the image texton library: sample 1 (felt), sample 4 (rough plastic), sample 6 (sandpaper), sample 10 (plaster-a), sample 12 (rough paper), sample 14 (roofing shingle), sample 16 (cork), sample 18 (rug-a), sample 20 (styrofoam), sample 22 (lambswool), sample 25 (quarry tile), sample 27 (insulation), sample 30 (plaster-b zoomed), sample 33 (slate-a), sample 35 (painted spheres), sample 41 (brick-b), sample 45 (concrete-a), sample 48 (brown bread), sample 50 (concrete-c), sample 59 (cracker-b).

for recognition. For the creation of the 3D texton library, a set of registered images from known viewing and illumination directions is required. In the training stage, the 3D texton method also requires registered images from known viewing and illumination directions. For the image texton method, a set of unregistered images from unknown viewing and illumination directions is sufficient in both the training stage and the creation of the image texton library. For recognition, the 3D texton method has a direct technique which requires a set (typically 20) registered images of the novel surface from known view/illumination. The 3D texton method also has an iterative Markov Chain Monte Carlo recognition technique requiring only a single image from unknown viewing and illumination direction. However the reported recognition rate is lower and the use of an iterative algorithm adds processing time and convergence issues. The image texton method uses a single image from unknown view/illumination without the need for an iterative algorithm.

2.2 Texton Histograms

The representation used for training is the image texton histogram, denoted by $H(l)$, as a discrete function of the set of labels l induced by the texton library. The global distribution of the local structural feature is encoded in the image texton histogram. Image texton histograms are obtained by filtering the texture images using the same filter bank as the one employed in creating the texton library. The resulting feature vectors are labeled by finding the closest vector in the image texton library. Due to the strong dependency of the appearance of image texture on the imaging conditions, the image texton histogram $H(l)$ is a function of the viewing and illumination direction. The dependency of texture histograms on viewing and illumination direction is also addressed in [3] which describes an analytical model of the bidirectional histogram of pixel intensity and in [14] where bidirectional intensity histograms are obtained via surface imaging simulations. However, intensity histograms do not sufficiently represent complex surface structure for the task of robust recognition. Texton histograms provide a richer description at the cost of requiring exemplars for building a representation instead of an analytical model. A particular 3D textured surface, originally represented by a collection of texture images, measured under different imaging conditions, is characterized by a set of image texton histograms indexed by the viewing and illumination direction. This representation encodes the variations of global spatial distribution with the imaging conditions, while being invariant over different texture instances of the same class.

The resultant model, namely a large collection of a high-dimensional vectors in the texton histogram space, is very complex and it raises the problem of compressing this representation to a lower-dimensional one while retaining the statistical properties of the image texton histograms. An analogous problem occurs in object recognition addressed in [10], where the representation is based on images of the object under various poses. This set of images is compressed via PCA to obtain a compact lower-dimensional representation. Similarly, for texture recognition we invoke PCA to reduce the complexity of the analysis space, which is spanned by the collection of texton histograms for a specific texture sample. Note that performing PCA on the set of texton histograms does not imply assumptions of Lambertian reflectance, considering that the eigenanalysis is not done directly on the reflectance values. We make the assumption that the image texton histograms can be projected onto an lower dimensional orthogonal basis while still preserving its descriptive properties and we empirically show through our recognition results that our assumption is reasonable.

A key issue in achieving a compact texture representation is identifying which subset of images in the full BTF is sufficient to characterize the texture class. For the CUReT database, there are 205 images per sample and we use a portion of these images for training. Instead of choosing this number randomly, we analyze the importance of individual images in the overall representation and choose a core subset of representative images. This process is termed manifold reduction and is described in detail in Section 3.

2.3 Classification

The classification of novel images is depicted as the final stage in Figure 2. Each novel image is from a different sample than those used for creation of the image texton library. In addition the imaging parameters (viewing and illumination directions) are different from those in the training set. Classification is achieved by filtering the novel image with the multiscale filter bank involved in creating the image texton histogram, projecting the resulting feature vectors in the image texton space, and labeling the feature vec-

Sample	θ_v	ϕ_v	θ_i	ϕ_i
polyester	1.209	-2.984	0.393	-1.571
terrycloth	0.196	3.142	0.982	0.000
leather	0.285	-2.377	1.371	-0.041
velvet	0.437	-2.701	0.196	-1.571
pebbles	0.196	3.142	0.982	0.000
frosted glass	0.393	3.142	0.785	0.000
aluminum foil	0.196	3.142	0.982	0.000
rough tile	0.810	-2.702	0.615	-0.685
rug-b	0.196	3.142	0.982	0.000
sponge	0.196	3.142	0.982	0.000
rabbit fur	0.196	3.142	0.982	0.000
limestone	0.196	3.142	0.982	0.000
brick-a	0.196	3.142	0.982	0.000
human skin	0.196	3.142	0.982	0.000
salt crystals	0.196	3.142	0.982	0.000
linen	0.628	-1.795	0.662	-1.136
stones	0.196	3.142	0.982	0.000
white bread	0.196	3.142	0.982	0.000
wood-a	0.717	-1.271	0.877	-0.767
tree bark	0.196	3.142	0.982	0.000

Table 1: Viewing and illumination directions ($\theta_v, \theta_i, \phi_v, \phi_i$, in radians) for the most significant texture images corresponding to each of the samples involved in recognition.

tors in respect to the texton library. The texton histogram is computed and this histogram is projected into the universal eigenspace, where each of the samples of interest is defined by a manifold. The sample corresponding to the manifold closest to the projected point is reported as the match.

3 Manifold Reduction

Assume the full BTF of a textured surface is sampled densely so that we have a large collection of texture images. Let us denote the number of texture images for a specific texture sample by N and for the CUReT database $N = 205$. As discussed, the set of image texton histograms is employed to compute a parametric eigenspace, onto which all image texton histograms are projected. We hypothesize that the projected points lie on a continuous manifold, which is a function of the viewing and illumination directions. Let us denote the manifold by $M(s)$, where parameter s is indexed by the imaging parameters $\{\theta_v, \phi_v, \theta_i, \phi_i\}$, the polar and azimuthal angles of viewing direction, and the polar and azimuthal angles of illumination direction, respectively.

Starting from this comprehensive parametric representation of BTF, we further hypothesize that each textured sample is sufficiently characterized by a core subset of texture images. To determine this representative subset of images for each of the samples to be analyzed, we design a quantitative evaluation method for the importance of individual images in the overall representation. There are two main steps involved in this evaluation method. The first step is the construction of the reference manifold $M_R(s)$ in the parametric eigenspace, i.e. the manifold which contains all N projected points. Note that in constructing the reference manifold an ordering of the points was necessary. The chosen ordering is consistent among all different texture samples, being determined in the 3D space of pose and illumination angles: θ_v, θ_i, and $\phi = |\phi_v - \phi_i|$. More specifically, the manifold path is constructed by first finding the diameter of the set of points in the 3D space of angles. Then the path is started at one of the extreme points, and the next point on the path is the closest unassigned point; the metric employed is Euclidean distance. The second step consists of sequential removal of individual points from the manifold, one at a time, with the exception of the end points. The reduced manifold which differs from the reference manifold by just one point, the ith, is denoted by $M_i(s)$. The resulting $N-1$ manifolds (a reference manifold $M_R(s)$, and $N-2$ reduced ones $M_i(s)$, having only $N-1$ knot points) are further analyzed by computing the distance between each of the reduced manifolds, and the reference manifold. In practice, a manifold is stored as a list of multidimensional vectors. The distance between two manifolds is determined by first computing the set of corresponding point-to-point distances, and then by averaging this set of distances. Based on the distribution of the distances between $M_R(s)$ and the set of $M_i(s)$, we determine an ordinal scale of the relative importance of individual points on the parametric manifold. That is, the least important point is the one for which its removal from the manifold generates a reduced one closest to the reference, while the most important point is the one for which its removal from the manifold generates a reduced one farthest from the reference. Using this ordering we compute a new set of $N-2$ reduced manifolds, obtained by consecutively eliminating from the reference manifold subsets consisting of i ordered points, $i = 1 \ldots N-2$. The reduced manifold which differs from the reference manifold by a set of i least significant points is denoted by $M^i(s)$. Again the distances d_i between the reference manifold $M_R(s)$ and the reduced manifolds $M^i(s)$, $i = 1 \ldots N-2$, are computed, and the distribution of these distances is considered in deciding which of the reduced manifolds is a compact representation of the texture sample, while still retaining the main structural characteristics. This decision is taken by imposing that the representative reduced manifold $M^\alpha(s)$ is at distance d_α to the reference manifold $M_R(s)$, where:

$$d_\alpha < c \cdot d_{max}, \quad c \in [0, 1] \tag{1}$$

and

$$d_{max} = \max d_i, \quad i = 1, \ldots, N-2 \tag{2}$$

This thresholding is arbitrary, though it is consistent across all texture samples. The ordering of the points is dependent on the characteristics of the fine-scale geometry of the texture sample and therefore varies from sample to sample. To determine which subset of the N images defines a core set for each sample, the union of the most important images per sample is taken.

4 Experiments and Results

Our recognition experiment uses the CUReT database [6], a collection of over 60 different real-world texture surfaces, each observed with 205 various combinations of viewing and illumination directions. We use Q = 20 samples for creating the image texton library and they are illustrated in Figure 3. For each of these samples, we consider

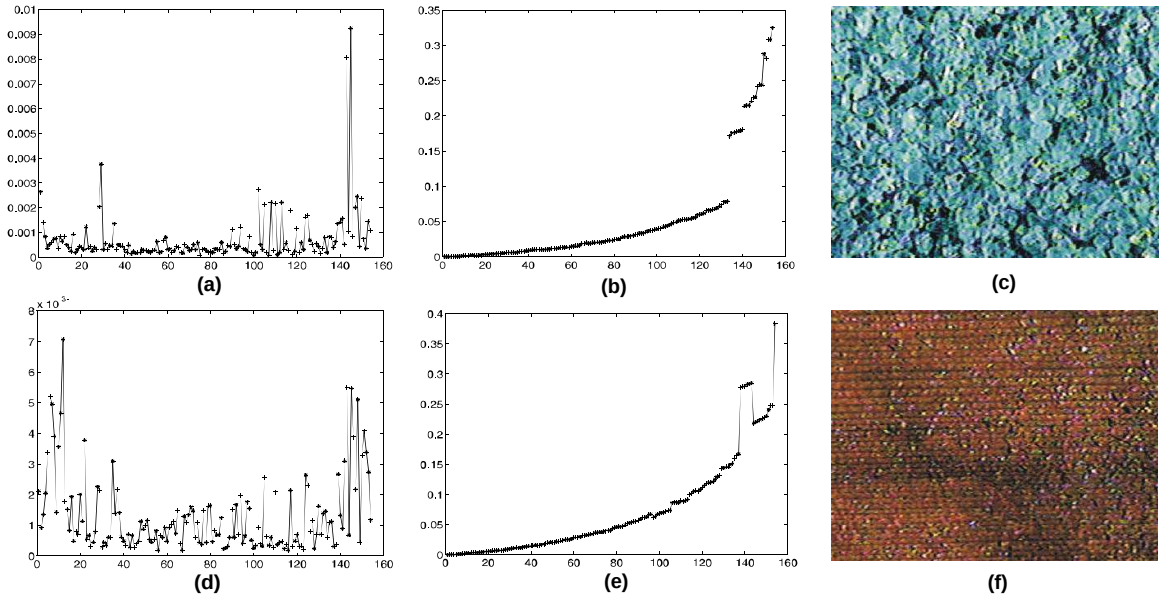

Figure 4: Samples 43 (salt crystals) and 54 (wood-a): (a,d) The distances between the reference manifold $M_R(s)$, based on N points, and the reduced manifolds $M_i(s)$, based on N-1 points, vs. the ordinal index i of the eliminated point; the reduced manifolds $M_i(s)$ are obtained by removing from the reference manifold one point at a time. (b,e) The distance between the reference manifold $M_R(s)$, based on N points, and the reduced manifolds $M^i(s)$, based on $N-i$ points, vs i. The reduced manifolds $M^i(s)$ are obtained after removing the first i least important points from the reference manifold. (c,f) Samples 43 and 54, illustrated by the corresponding most important texture images determined by reducing the manifold. It is important to notice how well the local structural configuration of the texture samples is captured in the texture images determined as the most important in representing the BTF.

q = 64 unregistered texture images of different viewing and illumination directions. The filter bank used to process the texture images consists of 48 oriented Gaussian derivative filters, center surround derivative filters, and low-pass Gaussian filters, on three different scales. For computational reasons we simplify the feature vector representation by grouping the filters on every scale. We define three different feature subspaces, each being spanned by the feature vectors obtained by aggregating the filter outputs corresponding to a scale. In order to obtain the image texton library, feature grouping via k-means algorithm is employed in all three feature subspace, and the result is the group of three image texton libraries, one per scale. For simplicity we will refer to the group of three texton libraries as a single library although the analysis is made in parallel for three different scales. Based on experimentation, we choose the size of the image texton library to be 100.

The resulting image texton library is used for computing the texton histograms of each of the $P = 20$ novel texture samples involved in recognition. As discussed, the set of samples for recognition and those used to create the texton library are disjoint. The novel samples are illustrated in Figure 5. For each novel sample we consider all available texture images, except those images captured under very tilted poses (polar angle of viewing direction θ_v greater than 70^o). For these very oblique views, the texture image becomes almost featureless and is therefore very difficult to classify. The resultant subset consists of $N = 156$ novel texture images for each sample, therefore in our work here we process for recognition a set of 3120 texture images.

For each of the texture samples we construct the parametric manifold into the corresponding eigenspace, by projecting all 156 image texton histograms, and by fitting a quadratic b-spline interpolation to the set of projected points.

As described in Section 3, we compute the compact representation for each of the novel samples. This reduced set of points on the manifold, encode the main structural characteristics of the BTF, and it is obtained by imposing the following constraint: the distance between the retained reduced manifold $M^\alpha(s)$ and the reference manifold $M_R(s)$ should not exceed $c = 30\%$ of the maximum distance between any of the reduced manifolds $M^i(s)$ and the reference $M_R(s)$. For two different samples, salt crystals (sample 43) and wood-a (sample 54), Figures 4 (a) and (d) illustrate the plots of distances between the reduced manifolds $M_i(s)$ and the reference $M_R(s)$ as a function of the index of the removed points i. It is interesting to observe how removal of different points on the manifold affects differently the representation in the eigenspace. That is, these two samples have very dissimilar local structure, therefore the most significant points do not coincide. Furthermore, from Figures 4 (b) and (e) it can be observed that the threshold of 30% of the maximum distance d_{max} between any of the reduced manifolds $M^i(s)$ and the reference $M_R(s)$ is reached sooner for sample 54 than for sample 43, and as a result the reduced representations of the two samples have different cardinalities. That is, for sample 43 a set of 23

Figure 5: The collection of 20 novel texture samples for which the recognition experiment is designed: sample 2 (polyester), sample 3 (terrycloth), sample 5 (leather), sample 7 (velvet), sample 8 (pebbles), sample 9 (frosted glass), sample 15 (aluminum foil), sample 17 (rough tile), sample 19 (rug-b), sample 21 (sponge), sample 24 (rabbit fur), sample 36 (limestone), sample 37 (brick-a), sample 39(human skin), sample 43 (salt crystals), sample 44 (linen), sample 47 (stones), sample 52 (white bread), sample 54 (wood-a), sample 58 (tree bark).

representative texture images is retained, while for sample 54, a number of 35 texture images are considered significant. As we expect, the number of significant images correlates well with the complexity of the observed surface structure. In Figure 4 (e), sample 54 presents a non-monotonic profile of distance d_i as a function of i, when i is large. This region of the plot corresponds to the reduced manifolds based on very few points, i.e. based on $N - i$ points, and as the most important texture images within the BTF are eliminated, the reduced manifolds become more unstable. In correlation with this, note that while sample 43 does not present this behavior, it also requires less texture images for the reduced representation, i.e. 23, in comparison with sample 54, which presents this behavior, and requires more significant images, i.e. 35, for constructing the compact representation. It is also interesting to notice how well the appearance of the two samples is captured in texture images illustrated in Figures 4 (c) and (f). These texture images correspond to the most significant images within the BTF of each sample, as found by our evaluation. This observation suggests the merit of using the method to identify the image or images which best conveys the surface texture.

Manifold reduction to determine the core image set is the key step to obtaining a compact representation. We experiment with various reduced BTF representations by varying the number of significant texture images considered for each texture sample. More specifically, the number of significant texture images included for each sample in the training set is varied from 11 to 1, decreasing with a step of 1. Note that when we form the global training set we take

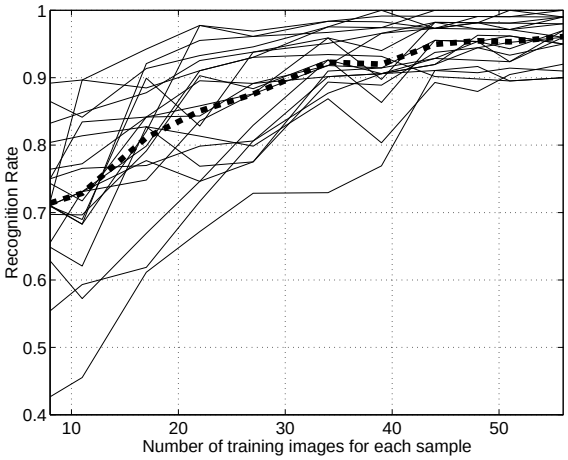

Figure 6: Recognition rates for individual texture samples (solid lines), and the global recognition rate (thick dashed line), as a function of the number of training images for each sample.

the union over all samples of all individual sets of significant images for each sample. For clarity consider the following example: suppose we assign numbers to the 205 images per sample in the CUReT database so that a particular number corresponds to a particular combination of imaging parameters; if image number 14 is in the set of significant images for any sample in the training set, then image number 14 from all the samples will be included in the training. By varying the cardinalities of the sets of most significant images per sample, we obtain corresponding training sets of various sizes, as listed in Table 2. Therefore we experiment with 11 variants of the training set. We also vary the set of texture images employed for testing, by considering each time the remaining images from the overall set of $N = 156$ images per sample.

The profile of the recognition rate as a function of the number of significant texture images used for training is depicted in Figure 6. By employing a number of 11 most significant texture images for each sample, we obtain a global training set of 1120 images, and consequently, a set of 2000 novel test images. An excellent global recognition rate of over 96% validates the efficiency of our approach. Also, for all samples involved in the recognition experiment we obtained individual recognition rates well above 90%. Moreover two of the samples reached 100% recognition rate, and other eight samples showed a recognition rate greater than 97%. These high recognition rates are impressive, especially considering that there is no need for alignment of texture images used for training the system, as well as there is no need for knowing the imaging conditions under which the images are captured. Also, during recognition, a single novel texture image with unknown viewing and illumination direction is employed, while the method itself is non-iterative.

By successively eliminating significant points from the individual sets of points, it can be observed that the recognition rate decreases. However, let us consider the minimal case, in which we take for each of the samples only one tex-

11	10	9	8	7	6	5	4	3	2	1
56	51	48	44	39	34	27	22	17	11	8

Table 2: The number of significant texture images considered for each of the texture samples (first row), and the corresponding cardinality p of the training set for each sample (second row).

ture image, the most significant. We list the imaging conditions for these images in Table 1. By taking the union over all samples, we obtain a global training set containing 160 images. In this case the testing is performed on 2960 texture images. Given this very restricted representation, due to the manner in which the training set is designed, we still obtain a global recognition rate well above 70%. Once more this demonstrates the utility of our representation.

However there are limitations, and one limiting factor is the breadth of the image texton library. Construction of texton library is based on a predefined set of texture samples, which need to be representative for a very large class of textured surfaces. In our experiments we found that two samples, sample 28 (crumpled paper) and sample 38 (ribbed paper), had local structural characteristics not well captured in the texton library, therefore the resultant models are not well discriminated.

5 Conclusion

In summary, we have developed a BTF-based texture representation that supports recognition or classification of textured surfaces invariant of illumination or viewing direction. In the test samples, surface appearance changes significantly with changes in viewing and illumination direction because the surface microgeometry introduces local occlusion, shadowing, foreshortening etc. Our representation is constructed in two stages. First an image texton library is constructed with a large variety of samples so that basic low level image features are well represented. The samples used to construct the image texton library are different from the samples used to recognition tests. The second stage is a training of surface types to be recognized. This training is accomplished by considering the texton histograms from a set of images obtained with different viewing and illumination directions. The dimensionality of the histogram space is reduced using PCA and a manifold is constructed for each sample. Texton histograms of novel samples are projected into the eigenspace of all samples and the closest manifold is determined using the SLAM software library [12]. The problem of determining which set of images to use for training is approached by quantitatively determining the effect of each image on the overall manifold to deduce a set that best represents the surface texture.

The methods described have important implications for computer vision. For texture recognition or classification, invariance to illumination or viewing direction is essential for robust performance. Our approach has the key advantages that a single image can be used for recognition without the need for iterative methods and the viewing/illumination direction of the novel images need not be known. Furthermore the images in the training set need not be registered. An additional impact of our work is the identification of a smaller subset of images from the CUReT database that can be used to sufficiently represent the texture class. The CUReT database contains over 200 images per sample and it is useful to determine which images can be considered redundant using a quantitative measure. Indeed this analysis can be used for BRDF/ BTF measurement planning in subsequent experiments.

The work described also has implications for computer graphics. Histograms of textural features have become an important building block for texture synthesis techniques. Compact representations of surface texture such as the one described here can be modified to support efficient and realistic rendering of natural objects with complex surface texture.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 0092491 and Grant No. 0085864.

References

[1] P. Brodatz. *Textures*. Dover, 1966.

[2] Oana G. Cula and Kristin J. Dana. Recognition methods for 3d textured surfaces. *Proceedings of SPIE Conference on Human Vision and Electronic Imaging VI*, 4299:pages 209–220, January 2001.

[3] K. J. Dana and S. K. Nayar. Histogram model for 3d textures. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 618–624, June 1998.

[4] K. J. Dana and S. K. Nayar. 3d textured surface modeling. *IEEE Workshop on the Integration of Appearance and Geometric Methods in Object Recognition*, pages 46–56, June 1999.

[5] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real world surfaces. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 151–157, June 1997.

[6] 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, January 1999.

[7] B. Julesz. Textons, the elements of texture perception and their interactions. *Nature*, (290):91–97, 1981.

[8] T. Leung and J. Malik. Recognizing surfaces using three-dimensional textons. *International Conference on Computer Vision*, 2:1010–1017, 1999.

[9] T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *Submitted to International Journal of Computer Vision*, 2000.

[10] H. Murase and S. K. Nayar. Visual learning and recognition of 3-d objects from appearance. *International Journal of Computer Vision*, pages 5–24, 1995.

[11] S. K. Nayar and H. Murase. Dimensionality of illumination in appearance matching. *IEEE Conf. on Robotics and Automation*, 1996.

[12] S. A. Nene, S. K. Nayar, and H. Murase. Slam: A software library for appearance matching. *Technical Report CUCS-019-94*, Proceedings of ARPA Image Understanding Workshop, November 1994.

[13] P. Suen and G. Healey. The analysis and recognition of real-world textures in three dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(5):491–503, May 2000.

[14] B. van Ginneken, J. J. Koenderink, and Kristin J. Dana. Texture histograms as a function of irradiation and viewing direction. *International Journal of Computer Vision*, 31(2-3):169–184, 1999.