Texture 2005 - In conjunction with ICCV
Pages 25-30

Statistical Estimation of Histogram Variation for Texture Classification

Robert E. Broadhurst
University of North Carolina
Department of Computer Science
Chapel Hill, NC 27599, USA
reb@cs.unc.edu

Abstract

I present a novel parametric approach for estimating the likelihood of homogeneously textured images. I propose that the dependence between pixel features is usefully captured by estimating the joint intra-class variation of their marginal distributions. To support this claim I build a single multivariate Gaussian distribution for each class that estimates the joint variation of several marginal, non-parametric, filter response histograms. I then generalize this framework to include marginal conditional distributions of pixel intensities for use with Strong-MRF models.

I demonstrate these methods on the Columbia-Utrecht database by classifying over 2800 images in all 61 texture classes. In a direct comparison with Varma & Zisserman (ECCV '02, CVPR '03) and Hayman (ECCV '04) this framework is found to be more accurate and efficient.

1 Introduction

This paper examines the general texture classification problem of material classification given images taken under varying and unknown viewing and illumination conditions. Algorithms for such tasks are characterized by four major components: features, the probability distribution representation, the probability distribution comparison, and the classifier. This work focuses on defining parametric Bayes classifiers with descriptive parameters that can be accurately estimated. The resulting method is used with both filters and Markov Random Fields (MRF) and is an extension of [7].

There are several classification algorithms related to this work. The MR8 classification algorithm of Varma and Zisserman uses a rotationally invariant filter bank and clustering to estimate the full joint probability distribution of filter responses [12]. Representative cluster centers define a texton dictionary, yielding a texton histogram representation for each image. The algorithm then uses the χ^2 distance and a 1-Nearest Neighbor (NN) classifier. Hayman extended the algorithm using Support Vector Machines [5]. Cula and Dana also construct a similar algorithm where principal component analysis (PCA) is applied to the texton histograms before using 1-NN with an unnormalized χ^2 distance [3].

Varma and Zisserman developed an alternative method based on MRFs [13]. MRFs model the conditional probability of a pixel's intensity given pixel intensities in a local neighborhood. This is estimated by first using textons to represent configurations of the neighboring intensities. Next, the distribution of intensities for each configuration is measured using each texton's assigned pixels. This variant uses the original χ^2 distance and 1-NN classifier.

Levina developed a classification algorithm based on histograms of marginal distributions [7, 8]. Her algorithm uses a filter bank, a 1-NN classifier, and a distance between two images set as the product of Mallows distances between corresponding marginal histograms.

In this paper, I extend Levina's filter bank based framework by using a Gaussian Bayes classifier instead of a 1-NN classifier. I use a multivariate Gaussian distribution to model the intra-class variability of each marginal histogram. This is accomplished by sensibly mapping histograms to points in Euclidean space so that PCA can then be applied. The mapping is chosen so that the Euclidean distance between two mapped histograms corresponds to their Mallows distance. In section 2, I describe several desirable properties of the Mallows distance and this space. Such a description is lacking for the Euclidean space assumed in [3, 5]. The authors use PCA and SVMs, respectively, in a space based on an unnormalized χ^2 distance. The methods in [7, 12, 13] use 1-NN classifiers, which must compute a distance from every test image to every training image. Gaussian Bayes classifiers are computationally efficient and, as shown in section 4, are more accurate than 1-NN classifiers.

In [3, 5, 12] the distribution of filter responses is estimated using a joint texton approach. I estimate the marginal distribution of each filter response using non-parametric histograms, which has several advantages. First, there is no

need to form a texton library, which limits generalizability and requires clustering in a high dimensional space. Second, marginal distributions can be very accurately estimated while joint distributions suffer from the curse of dimensionality. Although marginal distributions are less descriptive than their corresponding joint distribution, I partially capture the dependence among filter responses by estimating the joint intra-class variation of the marginals. This approach is shown in section 4.2 to increase the descriptive power of the marginals while maintaining their computational efficiency.

I apply this method not only to filter bank features as discussed above but also for features defined by Strong-MRF models [9]. Strong-MRF models approximate the joint probability of a neighborhood by assuming the conditional independence of the neighboring pixels given the center pixel. MRF models simplify texture classification by eliminating the need for filter bank design and response collection. Furthermore, MRF models require smaller support, allowing for more pixels with valid neighborhoods in an image. A possible disadvantage is that more features may be required complicating clustering in the joint conditional probability space. Strong-MRF models, however, avoid this complication by measuring the conditional probability of each neighboring pixel independently.

The remainder of the paper is organized as follows. Section 2 reviews the Mallows distance and describes properties of the corresponding Euclidean space. Section 3 describes my test methodology and database and gives 1-NN classification results. Section 4 extends this framework to Gaussian Bayes classification. Section 5 gives a generalization for images described by Strong-MRF models.

2 Constructing a Euclidean Space

In this section I construct a mapping that takes marginal histograms to points in Euclidean space. Such a mapping allows for the use of standard statistical tools to estimate intra-class histogram variation. Image classification requires the combination of the marginal results into an image description. This final step is classifier specific and is explored in section 3 for 1-NN classifiers and section 4 for Gaussian Bayes classifiers.

2.1 Mapping Marginal Distributions

In characterizing a mapping of a distribution to a point in Euclidean space, one should consider the distribution similarity measure corresponding to Euclidean distance. In this work the measure is the Mallows distance, a metric first used for texture classification by Levina [7]. Levina showed its equivalence (in many situations) to the Earth Mover's distance [8], used earlier by Rubner et al. for image retrieval [10, 11]. I use a marginal histogram representation similar to that used in [7] for 1-NN texture classification and in [1] for pixel intensity based image segmentation.

Intuitively, the Mallows distance measures the work required to change one distribution into another by moving probability mass. Therefore, this measure accounts for both the frequency and position of probability mass, unlike the χ^2 distance measure. This yields the desirable property that over-binning a histogram, or even using its empirical distribution, has no additional consequences other than measuring any noise present in the distribution estimate.

The Mallows distance between continuous one-dimensional distributions q and r, with cumulative distribution functions Q and R, respectively, is defined as

$$M_p(q,r) = \left(\int_0^1 |Q^{-1}(t) - R^{-1}(t)|^p dt \right)^{1/p}.$$

For example, consider two Gaussian distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$. For $p = 2$, their Mallows distance can be shown to equal $\sqrt{(\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)^2}$.

For discrete one-dimensional distributions, consider two distributions x and y represented by empirical distributions with n observations, or equi-count histograms with n bins and the average value of each bin stored. Considering these values in sorted order, x and y can be represented as vectors $\overline{x} = n^{-1/p} * (x_1, \ldots, x_n) = (x'_1, \ldots, x'_n)$ and $\overline{y} = n^{-1/p} * (y_1, \ldots, y_n) = (y'_1, \ldots, y'_n)$ with $x_1 \leq \ldots \leq x_n$ and $y_1 \leq \ldots \leq y_n$. The Mallows distance between x and y is then defined as the L_p vector difference between $\overline{x}$ and $\overline{y}$

$$M_p(x,y) = \left(\sum_{i=1}^{n} \|x'_i - y'_i\|^p \right)^{1/p}.$$

Therefore, this representation maps histograms to points in n-dimensional Euclidean space in which distances correspond to M_2 histogram distances.

This space contains several linear properties of interest. First, for histogram x with mean μ, $\overline{x} + n^{-\frac{1}{2}} \Delta\mu \overline{1}$ modifies x by only changing μ. Second, $\alpha\overline{x} - n^{-\frac{1}{2}}(\alpha - 1)\mu\overline{1}$ scales the standard deviation of x by α while leaving μ constant. Third, histograms of Gaussian distributions exist in a linear two-dimensional subspace. As for general distributions, one axis of this space represents a Gaussian's mean. The remaining orthogonal direction is linear in a Gaussian's standard deviation. Lastly, points in a convex portion of this space represent valid histograms. That is, a point $\overline{x}$ is a valid histogram if and only if $x_1 \leq \ldots \leq x_n$. Therefore, the mean of a set of histograms and any interpolated histogram will always be valid.

These properties help justify this representation of marginal distributions. In the next section I describe a specific classification problem before using this mapping in section 4 as a basis for Gaussian Bayes classification.

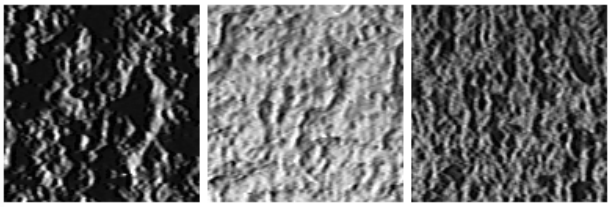

Figure 1. Three images from the "Plaster B" sample in the CUReT database illustrating the large intra-class variability.

Table 1. Classification accuracy using the MR8 filter bank, marginal histograms, the Mallows distance, and a 1-NN classifier.

# Bins	Independent Marginals	Joint Marginals
10	$93.63 \pm 1.05\%$	$95.35 \pm 0.95\%$
100	$95.14 \pm 0.90\%$	$96.35 \pm 0.82\%$
1000	$95.40 \pm 0.89\%$	$96.53 \pm 0.80\%$

3 Preliminaries

As in [5, 12, 13], this paper examines the classification of materials given images taken under varying and unknown viewing and illumination conditions. Section 3.1 further defines the test methodology, section 3.2 introduces the filter bank, and section 3.3 gives classification results using a 1-NN classifier. In section 4, I compare these results to Gaussian Bayes classification.

3.1 The CUReT Database

The CUReT database contains 61 texture classes consisting of materials imaged under 205 viewing and illumination conditions [4]. Each class contains images from one material that experience 3D effects such as specularities, interreflections, and shadowing, as shown in Figure 1. This large intra-class variability makes correct classification of the database a challenging task. The limitations of the database include a lack of significant scale change and limited in-plane rotation.

I follow the experimental setup of [5, 12, 13]. Of the 205 viewing and illumination configurations, 92 contain the largest minimum number of valid pixels across the samples. I use the cropped images of these 92 configurations supplied by Varma [12]. All results are averaged over 20 random splits of these images into equally sized test and training cases, yielding a total of $61 \times 46 = 2806$ training images and 2806 test images for each split. As a final preprocessing step, each image is converted to grey scale and processed to have zero mean and unit variance.

3.2 The MR8 Filter Bank

The MR8 filter bank consists of 38 filters and 8 filter responses [12]. There are two isotropic filters, a Gaussian and a Laplacian of a Gaussian (LOG), both at scale $\sigma = 10$. The 36 other filters include an edge (first derivative) filter at 6 orientations and 3 scales, and a bar (second derivative) filter at the same 6 orientations and 3 scales $(\sigma_x, \sigma_y) = \{(1, 3), (2, 6), (4, 12)\}$. Rotational invariance is achieved by storing only the maximum response over all orientations of a given filter type and scale. I normalize each filter response by the maximum attained over the training set.

3.3 1-NN Classification

Levina developed a classification framework using filter banks, marginal histograms, the Mallows distance, and 1-NN classification, which she demonstrated on the MeasTex and Brodatz databases [7]. In this section I implement a similar framework using the MR8 filter bank and compare it to Gaussian Bayes classification in section 4.

1-NN classifiers require a distance measure between two sets of marginal distributions. As in [7], I define this to be the product of the M_2 marginal distances described in section 2. The variation of marginal distributions can be measured jointly or independently. A joint 1-NN classifier measures the distance between a target image and all the training images as the distance between each set of marginals. The target image is then classified using the closest training image. For an independent 1-NN classifier, the minimum M_2 distance between each target marginal and each class is computed. The total distance to a class is defined as the product of each minimum marginal distance.

I represent each filter response distribution using equicount histograms containing 10, 100, or 1000 values. As shown in Table 1, the independent and joint 1-NN classifiers achieve a maximum accuracy of $95.40 \pm 0.89\%$ and $96.53 \pm 0.80\%$, respectively. The algorithm of Varma and Zisserman using the MR8 filter bank, a joint distribution estimate, and a 1-NN classifier achieves an accuracy of 96.93% and 97.43% with 610 textons and 2440 textons, respectively [12]. The SVM extension of this algorithm gives the best-known accuracy (98.46%) on this database [5]. While these algorithms are more accurate, they are also computationally more complex. In the next section, I improve upon these results and further reduce computational complexity by using a Gaussian Bayes classifier.

4 Gaussian Bayes Classification

I now improve upon the previous results by computing a parametric estimate of the variation of marginal non-parametric histograms using multivariate Gaussian distributions. Gaussian models stretch the Euclidean space of the mapped marginal histograms, enhancing the M_2 metric to account for the variability in the training set. I construct two Gaussian classifiers that model the variation of the marginal histograms independently and jointly in sections 4.1 and 4.2, respectively.

Table 2. Accuracy of the Gaussian Bayes classifier with independent or joint estimation of the MR8 marginal histograms.

# Bins	Independent Marginals	Joint Marginals
10	$94.34 \pm 0.94\%$	$99.20 \pm 0.37\%$
25	$95.16 \pm 0.83\%$	$99.22 \pm 0.34\%$
100	$95.65 \pm 0.83\%$	$99.03 \pm 0.38\%$
1000	$95.57 \pm 0.82\%$	$98.83 \pm 0.50\%$

4.1 Independent Marginal Variation

First, I consider the variation of the 8 marginal distributions defined by the MR8 filter bank to be mutually independent. For each texture class and filter response marginal I construct an independent multivariate Gaussian model. Each Gaussian model is trained using 46 histograms computed from the training set. Each histogram is represented using 10, 25, 100, or 1000 bins and is mapped to the Euclidean space of corresponding dimension. Note that 46 or more bins yields a high dimension low sample size (HDLS) situation. To accurately model this limited training set I use PCA to estimate a Gaussian distribution in a low dimensional subspace for each class and marginal. Since target images from other classes can have large projection errors, I measure the expected distance to the subspace by summing the remaining eigenvalues. For each class and marginal the final Gaussian model is of dimension equal to the number of eigenmodes plus one. For 10, 25, 100, and 1000 bins I use a small number of eigenmodes, 6, 7, 8, and 9, respectively, to define each subspace. For the 10 bin example, each class contains 8 independent 7-dimensional Gaussian models.

The likelihood of an image belonging to a class is defined as the product of the Gaussian likelihood of each marginal. During testing, an image is assigned to the class with the maximum likelihood of containing the image. This produces a Bayes classification of the image because the prior probability of each class is identical.

Table 2 shows the results of using this classifier with the marginal MR8 algorithm from the previous section. Compared to the independent 1-NN classifier, this classifier achieves a significant gain in speed along with a slight increase in accuracy. I discuss these results in the next section after defining the joint Gaussian classifier.

4.2 Joint Marginal Variation

Finally, I estimate a single Gaussian model per class that captures the joint variation of the feature marginals. For each image I combine into a single vector the reduced marginal representations computed in section 4.1. This representation maps images to points in Euclidean space where Euclidean distance corresponds to the L^2 norm of approximate M_2 marginal distances.

To construct a single class specific subspace, I concatenate each feature's subspace, disregarding the expected projection error, and apply PCA. Next, I record the expected projection error of the original histograms onto this final subspace by summing the 8 feature specific projection errors and the final projection error. For the following results I build a single 26-dimensional multivariate Gaussian distribution for each class, which measures the joint variation of a 25-dimensional subspace and the corresponding expected projection error. I use feature specific subspaces of dimension 15, except for the 10 bin case which uses all 10 available dimensions.

Gaussian classification using this model achieves an accuracy of 99.22% with 25 bins, as shown in Table 2. In addition, the classifier achieves a stable accuracy of over 99% for a large range of bin numbers and eigenmodes. The drop in classification accuracy for the 100 and 1000 bin representations is due to the HDLS problem of accurately estimating the corresponding multivariate Gaussian distributions. Hence, there is a tradeoff between accurate estimation of the marginal histograms and accurate estimation of the Gaussian models.

This algorithm is computationally efficient and it generates a compact representation of each image. For the classification problem being examined on the CUReT database, it is sufficient to capture only 80 values per image (10 per marginal), which can be further reduced to 25 values.

Thus far I have described four classifiers. With 10 bins per marginal, the independent 1-NN classifier, joint 1-NN classifier, independent Gaussian classifier, and joint Gaussian classifier achieve accuracies of 93.63%, 95.35%, 94.34%, and 99.20%, respectively. These results demonstrate the increase in accuracy of both Gaussian Bayes classification over 1-NN classification and of jointly modeling the variation of the marginal distributions over assuming their independence.

5 The Strong-MRF Framework

MRF models estimate probabilities of complete local neighborhoods of pixel values. To develop a classification algorithm using features defined by MRF models instead of filter banks, I need a simplification similar to that of marginal distributions of filter responses. Strong-MRF models provide such a simplification because they make additional Markovian assumptions equivalent to assuming the conditional independence of neighboring pixels given the center pixel's intensity. Paget describes Strong-MRF models in detail and demonstrates that they capture sufficient information for the synthesis of some natural textures [9]. Strong-MRF models are also related to gray level co-occurrence matrices since they both measure the co-occurrence of pairs of pixel intensities for several spatial relationships.

In section 5.1 I construct a mapping of the resulting marginal conditional distributions to Euclidean space. In section 5.2 I perform 1-NN and Gaussian classification and report results for neighborhoods of size 1×1, 3×3, 5×5, and 7×7. In section 5.3 I examine multi-scale neighborhoods.

5.1 Conditional Probability Representation

In this section I construct an estimate of the joint probability of local pixel neighborhoods consistent with Strong-MRF models and define a mapping of the resulting marginal conditional distributions to Euclidean space.

Let x_c be a pixel with intensity $I(x_c)$ and neighbors $y_1, y_2, \ldots, y_n$. I seek to estimate the joint probability

$$p(I(x_c), I(y_1), I(y_2), \ldots, I(y_n)).$$

Rewriting this as a conditional probability gives

$$p(I(x_c)) \cdot p(I(y_1), I(y_2), \ldots, I(y_n)|I(x_c)).$$

Next, assuming $y_1, y_2, \ldots, y_n$ are conditionally independent given x_c yields

$$p(I(x_c)) \cdot \prod_{i=1}^{n} p(I(y_i)|I(x_c)).$$

Finally, assuming x_c and $y_1, y_2, \ldots, y_n$ to be identically distributed and by the spatial symmetry of the neighborhood, the final joint probability estimate is rewritten as

$$p(I(x_c)) \cdot \prod_{i=1}^{n} p(I(x_c)|I(y_i)). \quad (1)$$

To construct a representation of each $p(I(x_c)|I(y_i))$ akin to the marginal representation, consider the conditional probability $p(X|Y)$. I represent $p(X|Y)$ as several equi-count histograms that estimate $p(X|Y)$ at a fixed number of intensity ranges of Y determined using equi-count bins. Each estimate is then measured with the same accuracy. To compute this representation, the values of Y are sorted and binned. Then, for each bin an equi-count histogram is formed using the corresponding X values. Thus, the final representation is an ordered set of equi-count histograms.

Table 3. Classification accuracy using $N \times N$ neighborhoods with a joint 1-NN classifier and a joint Gaussian Bayes classifier.

N	1	3	5	7
Joint 1-NN	61.45%	83.55%	88.23%	90.08%
Joint Gaussian	65.37%	92.79%	95.59%	96.04%

To define a distance between two conditional probabilities, consider $p(X|Y)$ and $p(Z|W)$. Let $y_1, y_2, \ldots, y_n$ and $w_1, w_2, \ldots, w_n$ be two sets of intensity ranges. I argue that such a distance measure only needs to compare $p(X|y_i)$ to $p(Z|w_i)$ for $i = 1, \ldots, n$ because X and Y, and Z and W, are identically distributed. Note that any differences between Y and W will have corresponding differences between X and Z. Thus, I ignore the actual intensity ranges of Y and W and define the distance between $p(X|y_i)$ and $p(Z|w_i)$ as their M_2 distance. Finally, I define the distance between $p(X|Y)$ and $p(Z|W)$ as the L^2 norm of the M_2 distances, which is a metric. This distance is the L^2 norm of the conditional probability representation taken as a vector so conditional probabilities can be considered as points in Euclidean space.

5.2 Local Neighborhoods

Applying Equation 1 to this conditional probability representation yields the following classification results. For 1-NN classification, I define the distance between two images to be the product of the distances between their marginal and conditional distributions. For Gaussian Bayes classification, I jointly model the variation of the distributions as in section 4.2.

The results in Table 3 show that the Gaussian classifier is significantly more accurate than the 1-NN classifier. The Gaussian classifier achieves an accuracy of $96.04 \pm 0.55\%$ with 7×7 neighborhoods. The results in Tables 3 and 4 use 10×10 histograms, where $j \times k$ bins for $p(X|Y)$ uses j bins to estimate $p(X|Y_i)$ at k configurations of Y. I use 18 eigenmodes with 1×1 neighborhoods and 40 feature specific and 30 joint eigenmodes otherwise. The MRF algorithm of Varma and Zisserman estimates the full conditional probability of a pixel's intensity given its local neigh-

Table 4. Classification accuracy using multi-scale 3×3 neighborhoods and a joint Gaussian Bayes classifier.

# Levels	1	2	3
Joint Gaussian	92.79%	97.57%	98.29%

borhood and achieves an accuracy of 95.87%, 97.22%, and 97.47% with 610 textons and 3×3, 5×5, and 7×7 neighborhoods, respectively [13]. Their best result of 98.03% with 2440 textons and a 7×7 neighborhood surpasses the results of the Gaussian classifier. These results, however, are improved in the next section using multi-scale neighborhoods.

5.3 Multi-scale Neighborhoods

The local neighborhood approach from section 5.2 uses information from a much smaller spatial extent around each pixel than filter based methods. While using larger neighborhoods should improve accuracy, there is a quadratic increase in the number of features with neighborhood size. Therefore, I use multi-scale neighborhoods to more compactly increase neighborhood extent and increase classification accuracy.

I define a pixel's multi-scale neighborhood to include the original 3×3 local neighborhood. Then, I use a Gaussian filter to generate 3×3 neighborhoods summarizing progressively larger spatial areas. Each level further blurs the image using a Gaussian filter with $\sigma = 1.4$.

As shown in Table 4, using a three level multi-scale neighborhood improves classification accuracy to $98.29 \pm 0.86\%$. The accuracy of this method demonstrates that my Strong-MRF model representation provides adequate information for classification and that filter banks are not necessary for this classification task. The filter bank method from section 4.2, however, is overall the most accurate and efficient, illustrating the effectiveness of a well designed filter bank.

6 Discussion and Conclusions

This paper presents several texture classification algorithms based on the statistical estimation of marginal histogram variation. The ability of these methods to describe the variation in the CUReT database is demonstrated by achieving a classification accuracy of over 99%. These results highlight the need to test this framework on more challenging problems using the CUReT database and others, including the KTH-TIPS database [5].

I plan to explore additional applications of this methodology to other texture analysis problems, including the synthesis of textures from the CUReT database at arbitrary viewing and illumination conditions, and the estimation of the viewing and illumination conditions of a given image.

In conclusion, this paper shows that algorithms based on joint marginal histograms, the Mallows distance, and Gaussian probability models are effective and efficient for texture classification. The algorithms presented here using the MR8 filter bank and MRF models achieve the best-known accuracy to date on the CUReT database.

Acknowledgements

The author is grateful to Stephen M. Pizer, J. Stephen Marron, and Sarang Joshi for discussions on histogram statistics and Gaussian parameter estimation. This work was partially supported by NIH grant P01 EB02779.

References

[1] R. E. Broadhurst, J. Stough, S. M. Pizer, and E. L. Chaney. Histogram statistics of local model-relative image regions. In *DSSCV*, 2005.

[2] O. G. Cula and K. J. Dana. Compact representation of bidirectional texture functions. In *CVPR*, 2001.

[3] O. G. Cula and K. J. Dana. 3d texture recognition using bidirectional feature histograms. *IJCV*, 59(1):33–60, 2004.

[4] K. Dana, B. van Ginneken, S. Nayar, and J. Koenderink. Reflectance and texture of real world surfaces. *ACM Trans. on Graphics*, 18(1):1–34, 1999.

[5] A. Hayman, B. Caputo, M. Fritz, and J. Eklundh. On the significance of real-world conditions for material classification. In *ECCV*, 2004.

[6] T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *IJCV*, 43(1):29–44, 2001.

[7] E. Levina. *Statistical Issues in Texture Analysis*. PhD thesis, 2002.

[8] E. Levina and P. Bickel. The earth movers distance is the mallows distance: Some insights from statistics. In *ICCV*, 2001.

[9] R. Paget. Strong markov random field model. *PAMI*, 26(3):408–413, 2004.

[10] J. Puzicha, Y. Rubner, C. Tomasi, and J. M. Buhmann. Empirical evaluation of dissimilarity measures for color and texture. In *CVPR*, 1999.

[11] Y. Rubner, C. Tomasi, and L. J. Guibas. A metric for distributions with applications to image databases. In *ICCV*, 1998.

[12] M. Varma and A. Zisserman. Classifying images of materials: achieving viewpoint and illumination independence. In *ECCV*, 2002.

[13] M. Varma and A. Zisserman. Texture classification: are filter banks necessary? In *CVPR*, 2003.