Nonparametric Multichannel Texture Description With Simple Spatial Operators

Timo Ojala and Matti Pietikäinen
Machine Vision and Media Processing Group
Infotech Oulu and Department of Electrical Engineering
University of Oulu, FIN-90570 Oulu, Finland
{skidi, mkp}@ee.oulu.fi

Abstract

A multichannel approach to texture description is proposed by approximating joint occurrences of multiple features with marginal distributions, as 1-D histograms, and combining similarity scores for 1-D histograms into an aggregate similarity score. A stepwise feature selection algorithm is used to choose the best feature combination in a particular dimension. In classification experiments with Brodatz textures and MeasTex test suites the proposed method performs favorably compared to GLCM, Gabor and GMRF features.

1. Introduction

Recently, we developed a set of new texture measures and conducted an extensive comparative study of texture measures with nonparametric classification based on feature distributions [1,2]. In this approach image texture is regarded as a distribution of a local property or local properties such as spatial pattern and pattern contrast, which are measured with simple spatial filters. These local texture operators transform the input, pixel values of a local neighborhood, into texture feature, which describes some property of the local image texture in this neighborhood. Four different types of operators were used: center-symmetric covariance measures [1], Local Binary Pattern operator [2], measures based on the gray level difference method [2] and Laws 3×3 masks [3].

In [1,2] only individual features or pairs of complementary features were used for texture discrimination. In most cases, a single measure cannot provide enough information about the local texture, tone or color. Better discrimination can be obtained by considering several features simultaneously. Also, we can consider a particular feature at multiple scales, by straightforwardly computing the desired feature for suitably symmetrical digital neighborhoods of any size, such as disks or boxes of odd or even size.

The main contribution of this paper is generalizing the texture description method to utilize multiple features and scales, by approximating joint occurrences of multiple features with marginal distributions. In the following chapters this generalized method is referred as the ‘Multichannel method.’

2. Multichannel texture description

To describe the texture of an image region the region is scanned with the local operator. The output values are accumulated into a discrete histogram with a fixed number of bins. For this purpose a mapping from the feature space to the bin index is needed [1]. Similarity between two histograms $S$ and $M$ is evaluated as a test of goodness-of-fit, which is measured with a nonparametric log-likelihood statistic, a two-way test of interaction or heterogeneity [4]:

$$G_2 = 2 \left( \sum_{S \times M} p_s \log p_s \right) - \left( \sum_{S \times M} p_s \log \left( \sum_{M} p_s \right) \right) - \left( \sum_{S \times M} p_s \log \left( \sum_{S} p_s \right) \right) \tag{1}$$

where $N$ is the number of bins and $p_i$ is the probability of bin $i$. The more alike histograms $S$ and $M$ are, the smaller is the value of the G-statistic $G_2$. The value of the test, when applied to a pair of histograms, reflects the probability that the histograms represent the same population. We employ the statistical test as a pseudo-metric, treating the value of the test as a distance measurement between the two histograms.

When several features and/or scales are used jointly, the question is how to combine these multiple channels into a
powerful description of image texture. One solution is to consider joint occurrences of the features, and approximate the joint distribution with a multidimensional histogram. As an example of this kind of approach, the gray level cooccurrence matrix estimates the joint gray level distribution for two gray levels located at a specified distance and angle. However, we can hardly expect to reliably estimate joint distributions for large numbers of features, even if the dimensions of individual features are small. Also, multidimensional histograms with large number of bins are very computationally intensive and consume very much memory.

An alternative is to use an approximation with marginal distributions and to employ each independent feature separately, as a one-dimensional histogram. Similarity scores are computed separately for each feature, and then combined into an aggregate similarity score. Based on the additivity property of the $G$-statistic (the results of several $G$-tests can be summed to yield a meaningful result), summing the individual scores seems natural, but for generality we explored the Minkowski norms

$$D_A(S, M) = \left[ \sum_{f=1}^{F} [D(S^f, M^f)]^R \right]^{1/R}$$

where $D_A$ is the aggregate score, $F$ is the number of individual features (dimensions), $D$ is the pseudo-metric used for measuring the similarity of histograms ($G_2$ in this paper), $S^f$ and $M^f$ are the respective sample and model histograms corresponding to feature $f$, and $R$ is the family parameter. Three values of $R$ were evaluated empirically: 1 (so-called ‘city-block’ distance), 2 (Euclidean distance), and $\infty$ (corresponds to taking the maximum of individual scores, the so-called ‘chessboard’ distance) of which $R=1$ resulted in the best performance. Puzicha et al. reached a similar conclusion in their study [9]. Hence the aggregate score $D_A$ is defined to be

$$D_A(S, M) = \sum_{f=1}^{F} D(S^f, M^f)$$

When an arbitrarily large number of features is available, it is important to choose nonredundant, uncorrelated measures. Generally, if the number of original features is large, a combinatorial explosion prevents us from performing an exhaustive search through all possible feature combinations. Feature selection was done with a stepwise search algorithm, which includes both forward and backward selection of features [5].

### 3. Experiments

#### 3.1. Brodatz textures

Fifteen fine-grained and homogeneous textures (Fig. 1) were taken from the Brodatz album [6]. Global gray-scale variations were removed with a Gaussian match [7]. Each 600x450 texture image was divided into 108 disjoint 50x50 samples, resulting in a classification problem involving 1620 samples. The data was partitioned into training and testing sets using the well-known leave-one-out method, i.e. while one particular sample was classified, the other 1619 samples served as models. In the classification process, the similarity between the sample and each model was measured with $D_A$ and the sample was assigned to the class to which the majority of the nine most similar models belonged (i.e. a 9-NN technique was applied).

Following 12 texture operators were considered in this problem. Three center-symmetric covariance measures: SAC, SRAC, SCOV; Local Binary Pattern operator LBP; four gray level difference features: DIFFX, DIFFY, DIFF2, DIFF4; four zero sum Laws 3x3 masks: L3E3, E3L3, L3S3, S3L3.

Given these operators, a stepwise search was performed with the chosen feature selection algorithm through all dimensions. Table 1 lists the error rate at each dimension. The smallest error rate of 0.19% was obtained with a set of five features: LBP, DIFF4, DIFF2, E3L3, and SAC. Inclusion of two highly correlated measures DIFF4 and DIFF2 is evidence of the suboptimal nature of the feature selection algorithm. Nevertheless, the result, with only three misclassified samples, is excellent, demonstrating the power of this approach. The classification accuracy saturated at 0.19%, and finally started deteriorating slowly, as the added features did not provide any useful information, but just increased the dimensionality (the ‘curse of dimensionality’). As the error rates for single features in [5] indicated, LBP was by far the best individual feature in this problem (1.98% error when the second best was SCOV with 16.91% error), which may skew the obtained result. In order to get a better idea about the gain achieved by using multiple features as marginal distributions, the test was repeated without LBP. 0.80% error was obtained with all 11 features.

For comparison, the same classification problem was tackled with the Gray Level Coocurrence Matrix (GLCM) features, the Gabor energy features, and the Gaussian Markov Random Field (GMRF) method. GLCM matrices were computed in the four principal directions with a spatial distance of one pixel. The number of gray levels was reduced from the original 256 to 32 with the equal probability quantizing algorithm. The 14 Haralick features were calculated from the four normalized cooccurrence matrices,
and the means and ranges of the four values comprised the set of 28 features used in the classification [8]. Gabor energy features were extracted with a filter bank of three different wavelengths (2, 4, and 8 pixels) and four different orientations (0, 45, 90, and 135 degrees). The width of the Gaussian window was set to wavelength/2 and the mask size was 5x5 pixels (different mask sizes between 3x3 and 21x21 pixels were attempted, of which 5x5 provided the best results). GMRF features were computed using the standard 4th-order symmetric mask (3rd, 4th, 5th and 6th order masks were attempted, of which the 4th-order mask gave best results). For classification purposes both a quadratic classifier and a 9-NN classifier were used, of which the quadratic classifier provided better results. The leave-one-out method was used for partitioning the data into training and testing sets, and a stepwise search was performed through all dimensions. GMRF features achieved an error rate of 1.67%, while error rates of 3.21% and 5.74% were obtained with GLCM and Gabor features, respectively. We see that the Multichannel method outperforms GMRF, GLCM and Gabor features in this problem. This is particularly impressive, because in contrast to GMRF and Gabor features the Multichannel method uses exclusively features computed at a single resolution (3x3).

**Table 1: Classification results for the Multichannel method, including and excluding LBP.**

<table>
<thead>
<tr>
<th>dimension</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>including LBP</td>
<td>1.98</td>
<td>0.49</td>
<td>0.43</td>
<td>0.31</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.25</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>excluding LBP</td>
<td>16.91</td>
<td>4.69</td>
<td>2.59</td>
<td>1.60</td>
<td>1.30</td>
<td>1.17</td>
<td>1.05</td>
<td>0.93</td>
<td>1.05</td>
<td>0.99</td>
<td>0.80</td>
<td>-</td>
</tr>
</tbody>
</table>

**Fig. 1. Fifteen Brodatz textures used in the experiments.**
3.2. MeasTex test suites

MeasTex is a quantitative framework for measuring the performance of a texture classification algorithm, maintained by CSSIP at the University of Queensland [10]. MeasTex provides a large database of homogeneous texture images, several test suites of texture classification problems, and implementations of major texture classification paradigms. The image database contains images from three different sources: Brodatz images, VisTex images and MeasTex images. VisTex images are from the Vision Texture database maintained by the Vision and Modelling group at the MIT Media Lab [11]. MeasTex images contain both natural and artificial textures. Natural textures include ‘grass’ textures (various types of grass), ‘materials’ textures (e.g. gravel and stone), and ‘surface’ textures (e.g. asphalt and concrete).

A test suite consists of a set of texture classification problems. A problem is specified to have a particular number of texture classes. For each class a set of training and validation images is provided, and generally the number of training and validation images in each class is at least 32. Most common image sizes are 32x32 and 64x64 pixels. A prior probability and a weight are assigned to each texture class. Based on the training images the texture classification algorithm builds a model of each texture class, and then each validation image is presented to the algorithm for classification. For each classification problem, the proportion of correctly classified validation images is defined to be the score of the problem (1.0 denotes perfect classification). The scores for individual problems are averaged to give the overall score for the test suite.

Following 21 operators were considered with the Multichannel method [12]. Seven center-symmetric covariance measures: SAC, SRAC, SCOV, SVR, VAR, BVAR, WVAR; Local Binary Pattern operator LBP; four gray level difference features: DIFFX, DIFFY, DIFF2, DIFF4; nine Laws 3x3 masks: L3E3, E3L3, L3S3, S3L3, E3S3, S3L3, E3E3, L3L3, S3S3.

However, no exhaustive search through all possible operator combinations was performed, but only a few combinations were tried, based on the individual operator performance in the given test suite. If quantization of the feature space was required, 32 bins were used. $G_2$ was used as a pseudo-metric, and the 9-NN technique was applied in classification.

Using four MeasTex test suites of natural textures, ‘Brodatz’, ‘grass’, ‘material’, and ‘VisTex’, the performance of the Multichannel approach was compared to that of three well-known paradigms: the Gray Level Co-occurrence Matrix (GLCM), Gabor energy, and Gaussian Markov Random Field (GMRF) methods. GLCM matrices were computed in the four principal directions with a spatial distance of one pixel. The images were requantized to have 32 gray levels. Rotation-invariance was achieved by averaging the features extracted from the four matrices. Gabor energy features were extracted with a filter bank of three different wavelengths (2, 4, and 8 pixels) and four different angles (0, 45, 90, and 135 degrees). The mask size was 17x17 pixels and the width of the Gaussian window was set to wavelength/2. GMRF features were computed using the standard 4th-order symmetric mask. A multivariate Gaussian classifier was used for classification.

<table>
<thead>
<tr>
<th>TEST SUITE</th>
<th>GLCM</th>
<th>Gabor</th>
<th>GMRF</th>
<th>Multichannel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brodatz (14)</td>
<td>0.8990</td>
<td>0.9419</td>
<td>0.9533</td>
<td>0.9930</td>
</tr>
<tr>
<td>grass (14)</td>
<td>0.9092</td>
<td>0.8797</td>
<td>0.9407</td>
<td>0.9751</td>
</tr>
<tr>
<td>material (22)</td>
<td>0.9597</td>
<td>0.9679</td>
<td>0.9670</td>
<td>0.9814</td>
</tr>
<tr>
<td>VisTex (24)</td>
<td>0.8062</td>
<td>0.8959</td>
<td>0.9223</td>
<td>0.9349</td>
</tr>
</tbody>
</table>

Scores for the four test suites are listed in Table 2. The number in parentheses corresponds to the number of individual problems in that test suite. The ‘material’ test suite contains both ‘materials’ textures and ‘surface’ textures. We see that the Multichannel method provides good results for all suites, ranking first in all of them.

4. Conclusions

This paper presented a simple but an efficient method for combining multiple features into a powerful description of image texture. In the classification experiments, the proposed method performed favorably compared to GLCM, Gabor and GMRF features, even though only features computed at a single resolution were used.

In this study similarity scores computed over 1-D histograms were combined into the aggregate similarity score. Instead of using only 1-D marginal distributions we could also consider joint 2-D distributions of pairs of complementary features, such as LBP/C and DIFFX/DIFFY, which seem to perform reliably even for relatively small sample sizes. Another aspect yet to be addressed is assigning varying weights to different features or feature pairs in the discrimination process, based on their individual performance on the problem in hand. In the proposed method all features contribute equally to the aggregate score, but obviously one would prefer better features to have larger effects and vice versa.
Acknowledgements

The financial support of the Academy of Finland and Tekes is gratefully acknowledged. The authors also wish to thank Professor Azriel Rosenfeld and David Harwood from the University of Maryland for their contributions.

References