PERGAMON

Pattern Recognition 33 (2000) 43–52

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/pattcog

Rotation-invariant texture classification using feature distributions

M. Pietikäinen*, T. Ojala, Z. Xu

Machine Vision and Media Processing Group, Department of Electrical Engineering, Infotech Oulu, University of Oulu, P.O. Box 4500, FIN-90401 Oulu, Finland

Received 25 July 1998; received in revised form 19 January 1999; accepted 19 January 1999

Abstract

A distribution-based classification approach and a set of recently developed texture measures are applied to rotation-invariant texture classification. The performance is compared to that obtained with the well-known circular-symmetric autoregressive random field (CSAR) model approach. A difficult classification problem of 15 different Brodatz textures and seven rotation angles is used in experiments. The results show much better performance for our approach than for the CSAR features. A detailed analysis of the confusion matrices and the rotation angles of misclassified samples produces several interesting observations about the classification problem and the features used in this study. © 1999 Published by Elsevier Science Ltd. All rights reserved.

Keywords: Texture analysis; Classification; Feature distribution; Rotation invariant; Performance evaluation

1. Introduction

Texture analysis is important in many applications of computer image analysis for classification, detection, or segmentation of images based on local spatial variations of intensity or color. Important applications include industrial and biomedical surface inspection; for example for defects and disease, ground classification and segmentation of satellite or aerial imagery, segmentation of textured regions in document analysis, and content-based access to image databases.

There are many applications for texture analysis in which rotation-invariance is important, but a problem is that many of the existing texture features are not invariant with respect to rotations. Some invariance for features derived from co-occurrence matrices or difference histograms can be obtained, for example, by simply averaging the matrices (histograms) or features computed for different angles, e.g. for 0, 45, 90 and 135°. Other early approaches proposed for rotation-invariant texture classification include the methods based on polarograms [1], generalized co-occurrence matrices [2] and texture anisotropy [3]. Kashyap and Khotanzad developed a method based on the circular symmetric autoregressive random field (CSAR) model for rotation-invariant texture classification [4]. Mao and Jain proposed a multivariate rotation-invariant simultaneous autoregressive model (RISAR) that is based on the CSAR model, and extended it to a multiresolution SAR model MR-RISAR [5]. A method for classification of rotated and scaled textures using Gaussian Markov random field models was introduced by Cohen et al. [6]. Approaches based on Gabor filtering have been proposed by, among others, Leung and Peterson [7], Porat and Zeevi [8], and Haley and Manjunath [9]. A steerable oriented pyramid was used to extract rotation invariant features by Greenspan et al. [10] and a covariance-based representation to transform neighborhood about each pixel into a set of invariant descriptors was proposed by Madiraju and Liu [11]. You and Cohen extended Laws' masks for

* Corresponding author. Tel.: +3588-553-2782; fax: +3588-553-2612.

E-mail address: mkp@ee.oulu.fi (M. Pietikäinen)

PII: S0031-3203(99)00032-1

rotation-invariant texture characterization in their "tuned" mask scheme [12].

Recently, we introduced new measures for texture classification based on center-symmetric auto-correlation and local binary patterns, using Kullback discrimination of sample and prototype distributions, and conducted an extensive comparative study of texture measures with classification based on feature distributions [13,14]. By using a standard set of test images, we showed that very good texture discrimination can be obtained by using distributions of simple texture measures, like absolute gray-level differences, local binary patterns or center-symmetric auto-correlation. The performance is usually further improved with the use of two-dimensional distributions of joint pairs of complementary features [14]. In experiments involving various applications, we have obtained very good results with this distribution-based classification approach, see e.g. Ref. [15]. Recently, we also extended our approach to unsupervised texture segmentation with excellent results [16]. An overview of our recent progress is presented in [17].

In our earlier experiments, the problems related to rotation invariance have not been considered. This paper investigates the efficiency of distribution-based classification, and our feature set in rotation-invariant texture classification. We are especially interested in seeing the performances for relatively small windows (64×64 and 32×32 pixels) required by many applications, whereas many of the existing approaches have been tested with larger windows. Texture measures based on center-symmetric autocorrelation, gray-level differences and a rotation-invariant version of the local binary pattern operator (LBPROT) are used in experiments. A simple method based on bilinear gray-level interpolation is used to improve rotation-invariance when extracting texture features in a discrete 3×3 neighborhood. The performance of LBPROT recently proposed by Ojala [18] is now experimentally evaluated for the first time. The use of joint pairs of features to increase the classification accuracy is also studied. The results for distribution-based classification are compared to those obtained with the well-known CSAR features.

2. Texture measures

2.1. Measures based on center-symmetric auto-correlation

In a recent study of Harwood et al. [13], a set of related measures was introduced, including two local center-symmetric auto-correlation measures, with linear (SAC) and rank-order versions (SRAC), together with a related covariance measure (SCOV). All of these are rotation-invariant robust measures and, apart from SCOV, they are locally gray-scale invariant. These measures are abstract measures of texture pattern and gray scale, providing highly discriminating information about the amount of local texture. A mathematical description of these measures computed for center-symmetric pairs of pixels in a 3×3 neighborhood (see Fig. 1) is presented in Eqs. (1)–(4). μ denotes the local mean and σ^2 the local variance in the equations.

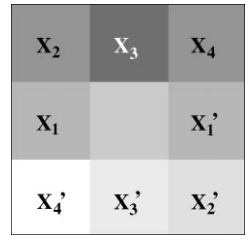

Fig. 1. 3×3 neighborhood with four center-symmetric pairs of pixels.

SCOV is a measure of the pattern correlation as well as the local pattern contrast. Since it is not 'normalized' in respect to local gray-scale variation, it provides more texture information than the normalized auto-correlation measures SAC and SRAC. SAC is an auto-correlation measure, a 'normalized', gray-scale invariant version of the texture covariance measure SCOV. SAC is invariant under linear gray-scale shifts such as correction by mean and standard deviation. It should also be noted that values of SAC are bound between -1 and 1.

$$\mathrm{SCOV} = \frac{1}{4}\sum_{i}^{4}(x_i - \mu)(x_i' - \mu), \tag{1}$$

$$\mathrm{SAC} = \frac{\mathrm{SCOV}}{\sigma^2}, \tag{2}$$

Texture statistics based directly on gray values of an image are sensitive to noise and monotonic shifts in the gray scale. With SRAC, the local rank order of the gray values is used instead of the gray values themselves. Hence, SRAC is invariant under any monotonic transformation including correction by mean and standard deviation and histogram equalization. The amount of auto-correlation in the ranked neighborhood is given by Spearman's rank correlation. It is defined for the $n \times n$ neighborhood with four center-symmetric pairs of pixels as (3), where m is n^2, and each t_i is the number of ties at rank r_i in the ranked neighborhood. The values of SRAC are bound between -1 and 1.

$$\mathrm{SRAC} = 1 - \frac{12\{\sum_i^4 (r_i - r_i')^2 + T_x\}}{m^3 - m}, \tag{3}$$

$$T_x = \frac{1}{12}\sum_{i}^{l}(t_i^3 - t_i). \tag{4}$$

The symmetric variance ratio (ratio between the within-pair and between-pair variances), SVR, is a statistic

equivalent to the auto-correlation measure SAC. SVR is also invariant under linear gray-scale shifts.

$$\mathrm{SVR} = \frac{\mathrm{WVAR}}{\mathrm{BVAR}}. \tag{5}$$

Additionally, the discrimination information provided by three local variance measures can be used. VAR (= BVAR + WVAR) and the two elements contributing to it are all measures of local gray-scale variation, very sensitive to noise and other local gray-scale transformations. The between-pair variance, BVAR, is mostly a measure of residual texture variance and usually it is a very small part of VAR. The majority of local variance is generally due to the within-pair variance WVAR. In our experiments, the classification results for the VAR measure are reported.

$$\mathrm{VAR} = \frac{1}{8}\sum_{i}^{4}(x_i^2 + x_i'^2) - \mu^2, \tag{6}$$

$$\mathrm{BVAR} = \frac{1}{16}\sum_{i}^{4}(x_i + x_i')^2 - \mu^2, \tag{7}$$

$$\mathrm{WVAR} = \frac{1}{16}\sum_{i}^{4}(x_i + x_i')^2. \tag{8}$$

2.2. Rotation-invariant local binary pattern

In the local binary pattern (LBP) texture operator we introduced [14], the original 3×3 neighborhood (Fig. 2a) is thresholded at the value of the center pixel. The values of the pixels in the thresholded neighborhood (Fig. 2b) are multiplied by the binomial weights given to the corresponding pixels (Fig. 2c). The result for this example is shown in Fig. 2d. Finally, the values of the eight pixels are summed to obtain the LBP number (169) of this texture unit. By definition LBP is invariant to any monotonic gray-scale transformation. The texture contents of an image region are characterized by the distribution of LBP.

LBP is not rotation invariant, which is undesirable in certain applications. It is possible to define rotation invariant versions of LBP; one solution is illustrated in Fig. 3 [18]. The binary values of the thresholded neighborhood (Fig. 3a) are mapped into an 8-bit word in clockwise or counterclockwise order (Fig. 3b). An arbitrary number of binary shifts is then made (Fig. 3c), until the word matches one of the 36 different patterns (Fig. 3d) of '0' and '1' an 8-bit word can form under rotation. The index of the matching pattern is used as the feature value, describing the rotation invariant LBP of this particular neighborhood.

2.3. Gray-level difference method

The method based on histograms of absolute differences between pairs of gray levels or of average gray levels has performed very well in some comparative studies and applications, see e.g. Ref. [14,15]. For any given displacement $\mathbf{d} = (\mathrm{d}x, \mathrm{d}y)$, where $\mathrm{d}x$ and $\mathrm{d}y$ are integers,

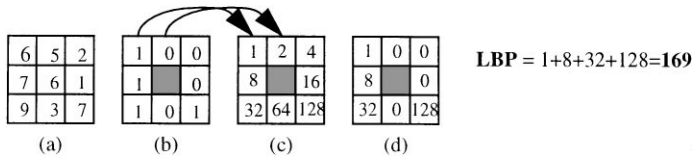

Fig. 2. Computation of local binary pattern (LBP).

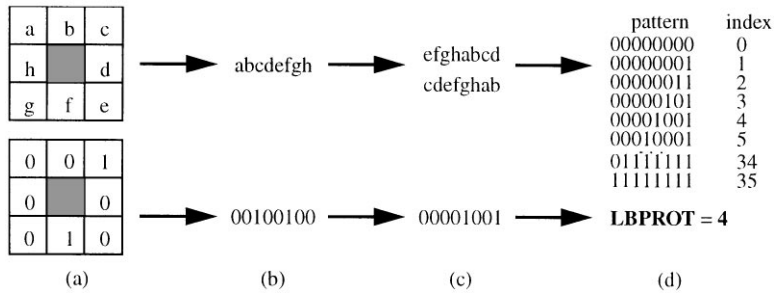

Fig. 3. Computation of LBPROT, rotation-invariant version of LBP.

let $f'(x, y) = |f(x, y) - f(x + \mathrm{d}x, y + \mathrm{d}y)|$. Let P' be the probability density function of f'. If the image has m gray levels, this has the form of an m-dimensional vector whose ith component is the probability that $f'(x, y)$ will have value i. P' can be easily computed by counting the number of times each value of $f'(x, y)$ occurs. For a small $\mathbf{d}$, the difference histograms will peak near zero, while for a larger $\mathbf{d}$ they are more spread out.

The rotation invariant feature DIFF4 is computed by accumulating, in the same one-dimensional histogram, the absolute gray-level differences in all four principal directions at the chosen displacement D. If $D = 1$, for example, the displacements $\mathbf{d} = (0, 1)$, $(1, 1)$, $(1, 0)$ and $(1, -1)$ are considered.

3. Classification based on feature distributions

Most of the approaches to texture classification quantify texture measures by single values (means, variances, etc.), which are then concatenated into a feature vector. In this way, much of the important information contained in the whole distributions of feature values is lost.

In this paper, a log-likelihood pseudo-metric, the G statistic, is used for comparing feature distributions in the classification process. The value of the G statistic indicates the probability that the two sample distributions come from the same population: the higher the value, the lower the probability that the two samples are from the same population. For a goodness-of-fit test the G statistic is

$$G = 2\sum_{i=1}^{n} s_i \log \frac{s_i}{m_i}, \tag{9}$$

where s and m are the sample and model distributions, n is the number of bins and s_i, m_i are the respective sample and model probabilities at bin i. In the experiments a texture class is represented by a number of model samples. When a particular test sample is being classified, the model samples are ordered according to the probability of them coming from the same population as the test sample. This probability is measured by a two-way test of interaction

$$\begin{aligned} G = 2\Bigg(&\left[\sum_{s,m}\sum_{i=1}^{n} f_i \log f_i\right] - \left[\sum_{s,m}\left(\sum_{i=1}^{n} f_i\right)\log\left(\sum_{i=1}^{n} f_i\right)\right] \\ &- \left[\sum_{i=1}^{n}\left(\sum_{s,m} f_i\right)\log\left(\sum_{s,m} f_i\right)\right] \\ &+ \left[\left(\sum_{s,m}\sum_{i=1}^{n} f_i\right)\log\left(\sum_{s,m}\sum_{i=1}^{n} f_i\right)\right]\Bigg), \end{aligned} \tag{10}$$

where f_i is the frequency at bin i. For a detailed derivation of the formula, see Ref. [19]. After the model samples have been ordered, the test sample is classified using the k-nearest-neighbor principle, i.e. the test sample is assigned to the class of the majority among its k nearest models. In our experiments, a value of 3 was used for k.

The feature distribution for each sample is obtained by scanning the texture image with the local texture operator. The distributions of local statistics are divided into histograms having a fixed number of bins; hence, the G tests for all pairings of a sample and a model have the same number of degrees of freedom. The feature space is quantized by adding together feature distributions for every single model image in a total distribution which is divided into 32 bins having an equal number of entries. Hence, the cut values of the bins of the histograms correspond to 3.125 (100/32) percentile of combined data. Deriving the cut values from the total distribution, and allocating every bin the same amount of the combined data guarantees that the highest resolution of the quantization is used where the number of entries is largest and vice versa. It should be noted that the quantization of feature space is only required for texture operators with a continuous-valued output. Output of discrete operators like LBP or LBPROT, where two successive values can have totally different meaning, does not require any further processing; operator outputs are just accumulated into a histogram.

To compare distributions of complementary feature pairs, metric G is extended in a straightforward manner to scan through the two-dimensional histograms. If quantization of the feature space is required, it is done separately for both features using the same approach as with single features.

4. Experimental design

4.1. Texture images

In the experiments, 15 classes of Brodatz [20] textures – pressed cork (D4), grass lawn (D9), herringbone weave (D16), woolen cloth (D19), french canvas (D21), calf leather (D24), beach sand (D29), pressed cork (D32), straw cloth (D53), handmade paper (D57), wood grain (D68), cotton canvas (D77), raffia (D84), pigskin (D92) and calf fur (D93) - were used. The original 600×450 images were globally gray-scale corrected by Gaussian match [21].

First, each image was rotated by 11° around its center using bicubic interpolation, in order to have a uniform behavior with respect to interpolation effects for both rotated and unrotated samples in the experiments. The operator included in the MATLAB Image Processing Toolbox was used [22]. In further discussion, we call these processed images reference images and they served as training data in the classification experiments. Then, the rotated images used for testing the texture classifier

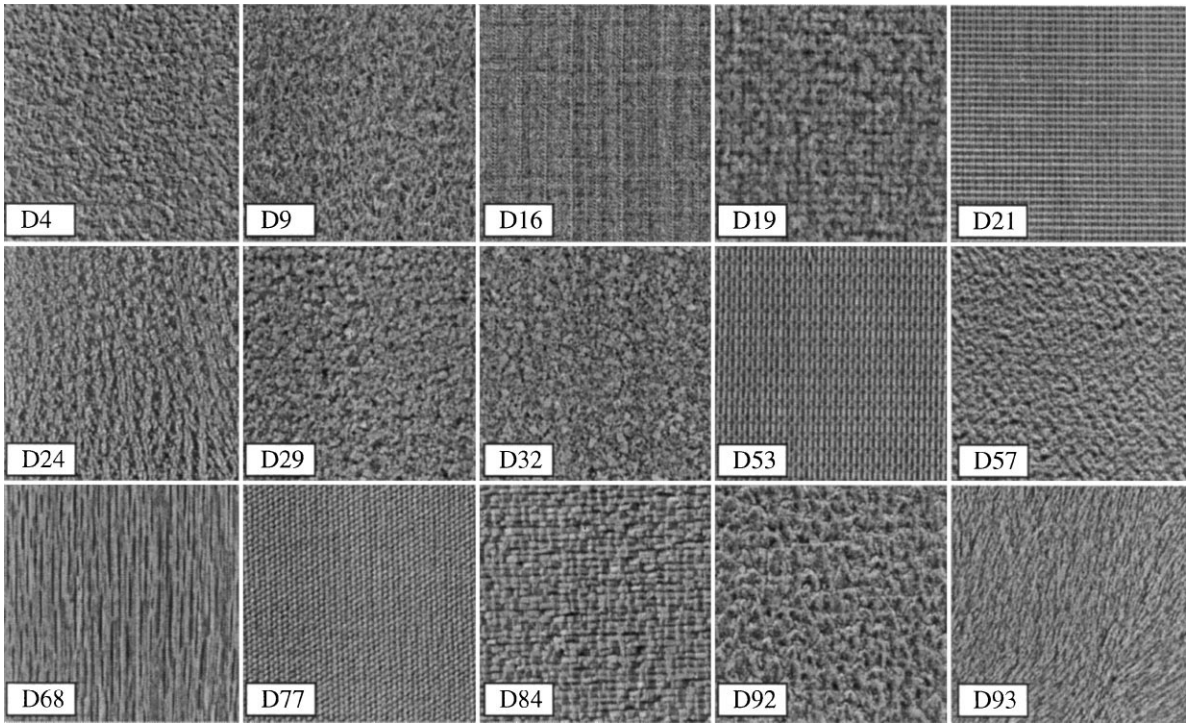

Fig. 4. Brodatz textures.

were generated by rotating each image counterclockwise around its center with the same bicubic interpolation method. We used the same set of six different rotation angles that was used by Kashyap and Khotanzad [4]: 30, 60, 90, 120, 150, and 200°. In other words, this is a true test of rotation-invariant texture classification, for the classifier 'sees' only instances of reference textures, and it is tested with instances of rotated textures it has not 'seen' before.

In order to analyze the effect of the window size, samples of two different sizes were used: 64×64 and 32×32 pixels. The samples were extracted inside a 256×256 rectangle located in the center of each texture image. Fig. 4 depicts the 256×256 images of each texture and Fig. 5 illustrates the extraction of sixteen 64×64 samples for a texture with a rotation of 30° (+ original rotation of 11°). Hence, with the window size of 64×64 pixels each texture class contained 16 reference samples for training the classifier and $6 \times 16 = 96$ rotated samples for testing it. Similarly, each texture class contained 64 training samples and $6 \times 64 = 384$ testing samples when the window size of 32×32 pixels was used. Even though the original texture images were globally corrected by Gaussian match, each individual sample image was also histogram equalized prior to feature extraction, to minimize the effect of possible local variations within the images.

4.2. Use of gray-scale interpolation and joint pairs of complementary features

A problem with computing rotation-invariant texture features in a local neighborhood is that the diagonal neighbors are farther from the center pixel than the horizontal and vertical neighbors, respectively. To reduce the effects of this on the classification performance, the pixel values for 'virtual' diagonal neighbors, located at the same distance from the center pixel than the horizontal and vertical neighbors, can be computed from the original pixel values by interpolation. In our experiments, we used a simple bilinear gray-scale interpolation method for this purpose.

In most cases, a single texture measure cannot provide enough information about the amount and spatial structure of local texture. Better discrimination of textures should be obtained by considering the joint occurrence of two or more features. In Ref. [14], we showed that the use of joint distributions of such pairs of features which provide complementary information about textures, usually improves the classification accuracy.

In this study, we perform experiments with various pairs of center-symmetric features and LBPROT. Our goal is not to find an optimal feature pair for this task, but to see how much improvement a joint feature pair can provide in the rotation-invariant classification problem.

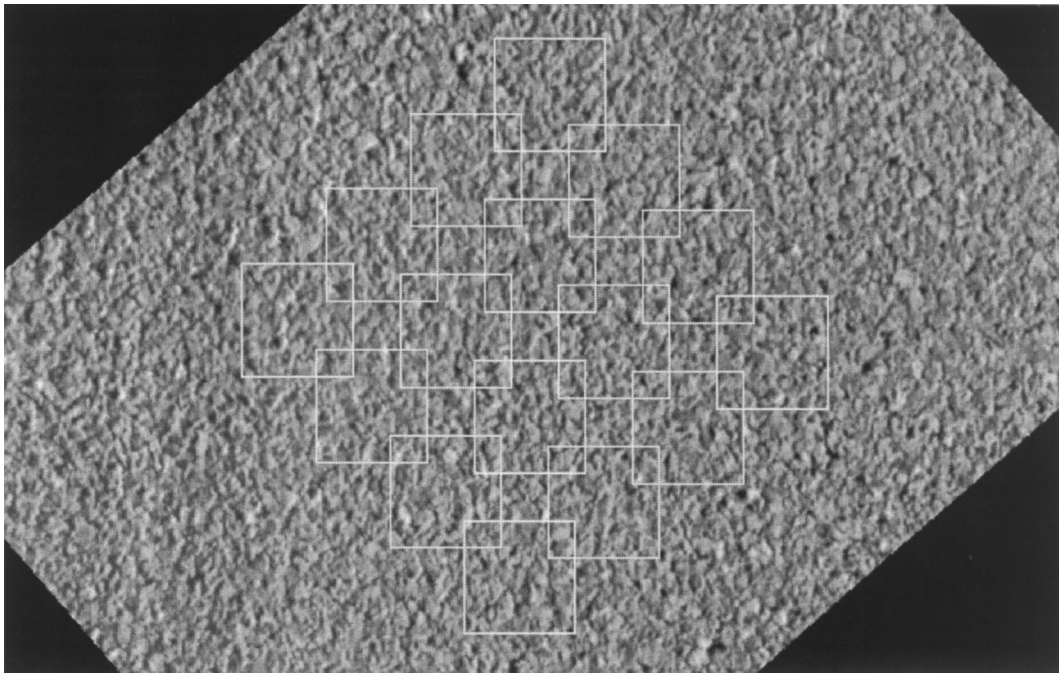

Fig. 5. Extraction of 64×64 samples from a rotated 600×450 image.

Table 1
Error rates (%) obtained with single features

Window size	Interpolation	DIFF4	LBPROT	SCOV	SAC	SRAC	SVR	VAR
64×64	no	28.5	38.5	27.6	37.1	41.2	36.9	24.0
	yes	27.2	39.2	16.3	32.4	48.1	31.8	14.2
32×32	no	44.9	50.5	39.0	53.6	48.4	53.9	40.7
	yes	39.6	47.7	35.1	50.0	59.4	50.3	36.4

5. Results and discussion

5.1. Single-feature performance

Table 1 shows the results of rotation-invariant classification for single features for window sizes of 64×64 and 32×32 pixels, without and with the interpolation of diagonal neighbors when features are extracted. In the case of 64×64 samples, VAR and SCOV features with interpolation provide the best error rates of 14.2 and 16.3%, respectively. DIFF4 also performs reasonably well with an error rate of 27.2%, whereas the worst results are obtained with the gray-scale invariant features (LBPROT, SVR, SAC, SRAC), indicating that information about local gray-scale contrast is useful in discriminating these Brodatz textures. We see that the use of gray-level interpolation usually improves the performance. Understandably, error rates for 32×32 samples are much higher, and the performance of, e.g. VAR deteriorates to 36.4%. The 32×32 samples contain only one fourth of the pixel data of the 64×64 samples, and consequently the feature histograms possess less discriminative power.

A closer examination of confusion matrices reveals that the center-symmetric features have most trouble in discriminating disordered textures. For example, in the case of 64×64 samples, textures D4, D9, D24, and D32 contribute almost 70% of the misclassified samples when feature VAR is used. Interestingly, LBPROT has no trouble separating D9 nor D32, missing only three of the 192 samples belonging to these two classes. This suggests that a better result could be obtained by carrying out classification in stages, selecting features which best discriminate among remaining alternatives. The local

gray-scale contrast seems to be particularly useful in separating textures D21, D57, and D77, for in these cases SCOV and VAR provide almost perfect results while their gray-scale invariant counterparts LBPROT, SVR, SAC, SRAC fail miserably. All features are able to discriminate textures D16 and D93, which both exhibit strong local orientation.

Similarly, a closer look at the rotation angles of misclassified samples reveals several interesting details. As expected, of the six rotation angles all features have the fewest misclassified samples at 90°. This attributes to the pseudo-rotation-invariant nature of our features. For example, both the center-symmetric features and DIFF4 are truly rotation-invariant only in rotations that are multiples of 45°, when computed in a 3×3 neighborhood, and in the chosen set of rotation angles 90° happens to be the only one of this type. A related observation, similarly attributed to the pseudo-rotation-invariant nature of the features, is that the results obtained at a particular rotation α and at its orthogonal counterpart ($\alpha+90°$) are very similar for all features. In other words, almost identical results are obtained at 30 and 120°, just like is the case with 60 and 150°. This suggests that the set of six rotation angles we adopted from Kashyap and Khotanzad [4] is suboptimal, at least when 3×3 operators are used. A more comprehensive test could be obtained by choosing a set of rotation angles that does not contain any two angles differing a multiple of 45°.

The rotation-invariance of the features can be improved by utilizing their generic nature. The features are not restricted to a 3×3 neighborhood, but they all can be generalized to scale. For example, the center-symmetric measures can be computed for suitably symmetrical digital neighborhoods of any size, such as disks or boxes of odd or even sizes. This allows for obtaining a finer quantization of orientations, for example, with a 5×5 box which contains eight center-symmetric pairs of pixels.

In earlier studies LBP, the rotation-variant ancestor of LBPROT, has proven to be very powerful in discriminating unrotated homogeneous Brodatz textures. However, in this study LBPROT provides fairly poor results with rotated Brodatz textures. By definition LBPROT is rotation-invariant only in digital domain, and consequently, of the six rotation angles, the interpretation of rotated binary patterns works properly only at 90°, where LBPROT provides clearly the best result of all features. In other angles, however, rotated binary patterns are obviously not mapped properly. This is particularly apparent in the case of strongly ordered textures (e.g. D21, D53, D68, and D77) where LBPROT provides perfect results at 90°, but fails completely at other rotation angles. This suggests that the current mapping of rotated binary patterns is probably far too strict, and could be relaxed by grouping together rotated patterns that have a specific Hamming distance, for example.

5.2. Results for joint pairs of features

The results obtained with joint pairs of features are presented in Table 2 and Table 3. The gray-scale interpolation was used in feature extraction and the histograms were quantized into 8×8 bins (36×8 for pairs including LBPROT). As expected, the use of feature pairs clearly improves the classification compared to the case of single features. We see that the feature pairs LBPROT/VAR and LBPROT/SCOV provide the best performance with error rates of 10.1 and 10.8% for 64×64 samples and 24.1 and 24.0% for 32×32 samples, respectively. Many other feature pairs also achieve error rates close to 10% for 64×64 samples, including SCOV/VAR, SCOV/SVR, SAC/SCOV and DIFF4/SAC. All these pairs have one thing in common: they include a feature that incorporates local gray-scale contrast. This emphasizes the importance of contrast in discriminating Brodatz textures. Consequently, the pairs of gray-scale invariant features fail.

In the previous section we made a remark about all features having the fewest misclassified samples at the rotation angle of 90°. This phenomenon is much stronger when joint pairs of features are used. For example, in the case of LBPROT/VAR only 5 of the 145 misclassified samples occur at 90°, while each other rotation angle contributes at least 24 misclassified samples.

Interestingly, even though LBPROT does fairly poorly by itself, LBPROT combined with a gray-scale variant

Table 2
Error rates (%) obtained with pairs of features for 64×64 samples

	LBPROT	SCOV	SAC	SRAC	SVR	VAR
DIFF4	17.6	12.8	12.6	30.6	12.9	16.3
LBPROT		10.8	29.2	38.1	29.2	10.1
SCOV			11.6	31.7	11.5	10.6
SAC				43.9	32.0	14.6
SRAC					44.2	23.4
SVR						14.4

Table 3
Error rates (%) obtained with pairs of features for 32×32 samples

	LBPROT	SCOV	SAC	SRAC	SVR	VAR
DIFF4	30.2	25.5	28.7	40.0	28.5	30.7
LBPROT		24.0	41.2	45.5	41.0	24.1
SCOV			28.7	38.1	28.6	29.8
SAC				54.5	50.9	28.4
SRAC					54.2	35.2
SVR						28.4

feature (SCOV or VAR) provides the best results of all pairings. This indicates that the combination of just crude pattern shape/code and pattern contrast is a useful description of rotated Brodatz textures. The shortcomings of LBPROT are still apparent, though, for in the case of 64×64 samples, strongly ordered texture D21 contributes 80 of the 145 samples misclassified by the LBPROT/VAR pair. These 80 misclassified D21 samples, which all are incidentally assigned to class D9, correspond to samples of all other rotation angles but those of 90° which are classified correctly. Obviously, a significant improvement in the classification accuracy can be expected, when the shapes of rotated binary patterns are described effectively.

6. Quantitative comparison to circular-symmetric autoregressive random field (CSAR) model

6.1. CSAR features

For comparison purposes we used the circular-symmetric auto-regressive random field (CSAR) model which was proposed for rotation invariant texture classification by Kashyap and Khotanzad [4]. The general expression of the model is described as

$$y(s) = \alpha \sum_{r \in N_c} g_r y(s \oplus r) + \sqrt{\beta} v(s), \tag{11}$$

where $\{y(s),\ s \in \Omega, \Omega = (0 \leqslant s_1, s_2 \leqslant M-1)\}$ is a set of pixel intensity values of a given $M \times M$ digitized discrete image, s are the pixel coordinates, N_c is the neighborhood pixel set, r are the neighborhood pixel coordinates, $v(s)$ is a correlated sequence with zero mean and unit variance, α and β are the coefficients of the CSAR model, and $\oplus$ denotes modulo M addition. The model from this equation yields two parameters, α and β. Parameter α measures a certain isotropic property of the image and β a roughness property. The rotation invariant characteristics of α and β are contributed by the choice of the interpolated circular neighborhood of image pixels. There is also a third parameter ζ which is a measure of directionality and is determined by fitting other SAR models over the image [4].

6.2. Classification procedure

Texture classification is performed by extracting parameters α, β and ζ from the sample images, and feeding them to a feature vector classifier. In other words, a texture sample is represented by three numerical CSAR features, in contrast to our approach where features are described by histograms of 32 bins. The texture classifier is trained with the features of the reference samples, and tested with the rotated samples. Both a multivariate Gaussian (quadratic) classifier and a 3-NN (nearest neighbor) classifier were used. When the 3-NN classifier was used, the features were first normalized into the range 0–1 by dividing each feature with its maximum value over the entire training data.

6.3. Results and discussion

Since the quadratic classifier provided slightly better results than the 3-NN classifier, the discussion is based on the results obtained with the quadratic classifier (Table 4). Even though none of the individual features is very powerful by itself (the best feature (β) provides error rates of 51.9 and 60.0% for the 64×64 and 32×32 samples, respectively), the three features combined offer a reasonable texture discrimination with error rates of 15.3 and 29.3%, respectively. However, we see that the distribution-based classification with single features (e.g. VAR and SCOV) performs about as well as the three CSAR features combined in the case of 64×64 samples, and several pairs of joint features provide better performance than the combined CSAR features for both the 64×64 and 32×32 samples.

Confusion matrices reveal that the CSAR features have difficulties in separating textures D4, D24, D57, and D68. When all three features are used, these four classes contribute almost 80% of the 221 misclassified 64×64 samples. The confusion is particularly severe between classes D4 (pressed cork) and D24 (pressed calf leather). Examination of the rotation angles of the misclassified samples verifies the observation of the pseudo-rotation-invariant nature of 3×3 operators made in Section 5.1, for again by far the fewest misclassified samples occur at 90° (only 10, while each other rotation angle contributes at least 31), and again the results obtained at two orthogonal rotation angles are almost identical.

Table 4
Error rates (%) obtained with the CSAR features

Window size	α	β	ζ	$\alpha+\beta$	$\alpha+\zeta$	$\beta+\zeta$	$\alpha+\beta+\zeta$
64×64	56.4	51.9	68.5	24.9	34.2	27.6	15.3
32×32	67.0	60.0	72.5	44.5	48.2	37.0	29.3

7. Conclusion

In this paper, a distribution-based classification approach and a set of texture measures based on center-symmetric autocorrelation and local binary patterns were applied to rotation-invariant texture classification. The performance of the proposed approach was compared to that of circular-symmetric autoregressive random field (CSAR) model with a difficult classification problem involving 15 different Brodatz textures and seven rotation angles. The error rates of the best single features (SCOV, VAR) were comparable to those obtained with the three CSAR features combined, and better results were achieved with distributions of joint pairs of features.

It was also shown that the rotation invariance of texture measures in a discrete 3×3 neighborhood can be improved by gray-level interpolation. The experimental results also emphasize the importance of local pattern contrast in discriminating Brodatz textures, for even though the samples were corrected against global gray-scale variations, the features measuring local gray-scale variations clearly outperformed their gray-scale invariant counterparts. Another observation, verified both with our operators and with the CSAR features, was the pseudo-rotation-invariant nature of the features which should be taken into account when designing new operators and new studies on rotation-invariant texture classification. The shortcomings of LBPROT, the rotation-invariant version of the powerful LBP operator, were exposed, and a significant improvement in classification accuracy can be expected, once the shapes of rotated binary patterns are described effectively.

Acknowledgements

The financial support provided by the Academy of Finland and Technology Development Center is gratefully acknowledged. The authors also wish to thank the anonymous referee for his helpful comments.

References

[1] L.S. Davis, Polarograms: a new tool for texture analysis, Pattern Recognition 13 (1981) 219–223.

[2] L.S. Davis, S.A. Johns, J.K. Aggarwal, Texture analysis using generalized co-occurrence matrices, IEEE Trans. Pattern Anal. and Mach. Intell. 1 (1979) 251–259.

[3] P. Chetverikov, Experiments in the rotation-invariant texture discrimination using anisotropy features, in: Proceedings of the 6th International Conference on Pattern Recognition, Munich, Germany, 1982, pp. 1071–1073.

[4] R.L. Kashyap, A. Khotanzad, A model-based method for rotation invariant texture classification, IEEE Trans. Pattern Anal. Mach. Intell. 8 (1986) 472–481.

[5] J. Mao, A.K. Jain, Texture classification and segmentation using multiresolution simultaneous autoregressive models, Pattern Recognition 25 (1992) 173–188.

[6] F.S. Cohen, Z. Fan, M.A. Patel, Classification of rotated and scale textured images using Gaussian Markov random field models, IEEE Trans. Pattern Anal. and Mach. Intell. 13 (1991) 192–202.

[7] M.M. Leung, A.M. Peterson, Scale and rotation invariant texture classification, in: Proceedings of the 26th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, 1992.

[8] M. Porat, Y. Zeevi, Localized texture processing in vision: analysis and synthesis in the Gaborian space, IEEE Trans. Biomed. Engng. 36 (1989) 115–129.

[9] G.M. Haley, B.S. Manjunath, Rotation-invariant texture classification using modified Gabor filters, in: Proceedings of the IEEE Conference on Image Processing, Austin, TX, 1994, pp. 655–659.

[10] H. Greenspan, S. Belongie, R. Goodman, P. Perona, Rotation invariant texture recognition using a steerable pyramid, in: Proceedings of the 12th International Conference on Pattern Recognition, vol. 2, Jerusalem, Israel, 1994, pp. 162–167.

[11] S.V.R. Madiraju, C.C. Liu, Rotation invariant texture classification using covariance, in: Proceedings of the IEEE Conference on Image Processing, vol. 1, Washington, DC, 1995, pp. 262–265.

[12] J. You, H.A. Cohen, Classification and segmentation of rotated and scaled textured images using texture "tuned" masks, Pattern Recognition 26 (1993) 245–258.

[13] D. Harwood, T. Ojala, M. Pietikäinen, S. Kelman, L.S. Davis, Texture classification by center-symmetric autocorrelation, using Kullback discrimination of distributions, Pattern Recognition Lett. 16 (1995) 1–10.

[14] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on feature distributions, Pattern Recognition 29 (1996) 51–59.

[15] T. Ojala, M. Pietikäinen, J. Nisula, Determining composition of grain mixtures by texture classification based on feature distributions, Int. J. Pattern Recognition Artificial Intell. 10 (1996) 73–82.

[16] T. Ojala, M. Pietikäinen, Unsupervised texture segmentation using feature distributions, Pattern Recognition 32 (1999) 477–486.

[17] M. Pietikäinen, T. Ojala, O. Silven, Approaches to texture-based classification, segmentation and surface inspection, in: C.H. Chen, L.F. Pau, P.S.P. Wang (Eds.), Handbook of Pattern Recognition and Computer Vision, second ed., World Scientific, Singapore, 1998 (chapter 4.2).

[18] T. Ojala, Multichannel approach to texture description with feature distributions, Technical Report CAR-TR-846, Center for Automation Research, University of Maryland, 1996.

[19] R.R. Sokal, F.J. Rohlf, Introduction to Biostatistics, second ed., W.H. Freeman, New York, 1987.

[20] P. Brodatz, Textures: A Photographic Album for Artists and Designers, Dover Publications, New York, 1966.

[21] J.M. Carstensen, Cooccurrence feature performance in texture classification, in: Proceedings of the 8th Scandinavian Conference on Image Analysis, vol. 2, Tromsö, Norway, 1993, pp. 831–838.

[22] Image Processing Tooolbox for use with MATLAB. The MathWorks, Inc. (1984–1998).

About the Author—MATTI PIETIKÄINEN received his Doctor of Technology degree in Electrical Engineering from the University of Oulu, Finalnd, in 1982. Currently he is Professor of Information Technology, Scientific Director of Infotech Oulu – a center for information technology research, and Head of Machine Vision and Media Processing Group, at the University of Oulu. From 1980 to 1981 and from 1984 to 1985 he was visiting the Computer Vision Laboratory at the University of Maryland, USA. His research interests include machine vision, document analysis, and their applications. He has authored about 100 papers in journals, books and conferences. He is the editor (with L.F. Pau) of the book "Machine Vision for Advanced Production", published by World Scientific in 1996. Prof. Pietikäinen is Fellow of international Association for Pattern Recognition (IAPR) and Senior Member of IEEE, and serves as Member of the Governing Board of IAPR and was Chairman of IAPR Education Committee. He also serves on program committees of several international conferences.

About the Author—ZELIN XU received the M.S. degree in Optics from the Changchun Institute of Optics and Fine Mechanics, China, in 1984, respectively. Currently he is a graduate student at the Department of Electrical Engineering of the university of Oulu, Finland. His research interest include texture analysis and visual inspection.

About the Author—TIMO OJALA received the M.S. degree in Electrical Engineering with honors from the University of Oulu, Finland, in 1992, and his Dr. Tech. degree from the same university in 1997. He is a member of Machine Vision and Media Processing Group at the University of Oulu. From 1996 to 1997 he was visiting the University of Maryland Institute for Advanced Computer Studies (UMIACS). His research interests include pattern recognition, texture analysis and object-oriented software design.