Systematic Feature Extraction Using Mask Patterns for High Precision Image Analysis

Takahiro TOYODA[†] and Osamu HASEGAWA[††,†††]
[†] Interdisciplinary Graduate School of Science and Engineering
[††] Imaging Science and Engineering Laboratory
Tokyo Institute of Technology
4259 Nagatsuta-cho, Midori-ku, Yokohama, 226-8503, JAPAN
[†††] PRESTO, JST
toyoda@isl.titech.ac.jp

abstract

We propose a texture feature extraction method using mask patterns of several sizes. A mask of each size has 223 patterns. A maximum of 233 features, including the 10 additional features, are extracted from each mask. We construct multi-resolution features by combining the features extracted from masks of different sizes. The calculated feature value corresponds to the power spectrum of a mask pattern within the image. This method is equivalent to expressing the basis functions in frequency analysis by the mask patterns. Texture classification experiments demonstrate that our proposed features outperformed other features such as Gabor. They also worked well in a test where the images to be classified were captured under different illuminants from training images. These results indicate a high practicality of the proposed method. This paper also shows that applying a feature selection method to the proposed features provides high recognition rates with a small number of features.

1 Introduction

Texture analysis is important in many applications like medical imaging and remote sensing. It is also used for segmentation to extract target objects from an image. The edge detection algorithm, by calculating the gradient, can also be used for segmentation. However, this algorithm does not always produce closed target regions. Moreover, it has difficulty in distinguishing the edges inside the region and on the boundary. On the other hand, obtained boundaries are always closed and each region is characterized by value when texture analysis is used.

In most image analyses, including those of robot vision and road sign recognition systems, which are both researched actively, the first procedure of image processing is the extraction of target objects. This step should be performed with high accuracy because its results affect the subsequent analytical procedures. This aspect indicates the greater importance of texture analysis and great demand for high-precision methods.

Texture features that are used for analysis have a great influence on performance. Analytical methods are usually divided into two approaches: structural approaches and statistical approaches. Structural approaches are applied to regular textures. They represent such textures by their elements and the arrangements of those elements. In contrast, statistical approaches are used for analysis of textures that have fine and irregular structures. They characterize such structures by the distribution of intensity. Statistical approaches include methods based on local features such as Higher order local autocorrelations (HLAC), Gaussian Markov Random Field (GMRF), and Local Binary Pattern (LBP), in addition to methods based on frequency analysis like Fourier transform and Gabor filters.

This paper proposes a statistical feature extraction method based on local regions. It uses mask patterns of various sizes and configurations. These features are shift-invariant (they can classify the texture images regardless of their positions). The calculated feature values represent a power spectrum of each mask pattern within the image.

To yield more effective features, we apply a basic feature selection method to the proposed features. Experimental results described in this paper demonstrate that this approach raises recognition accuracy while reducing the number of features markedly. The selection procedures are performed by the system automatically without a tedious manual operation.

The proposed method can calculate the feature values simply. Furthermore, hardware implementation is easy. Moreover, it is robust against illumination changes. All of these features underscore the practical effectiveness of this proposed method.

2 Features

2.1 Mask Patterns

We propose a texture feature extraction method using mask patterns. The approach of creating various mask patterns is considered to enhance the discrimination power of the features. Unless patterns are created appropriately, however, the number of obtained patterns will be enormous. In this study, we first restrict the size of the mask within a 3×3 pixel region, the center of which is the observing point and the eight neighboring pixels are the

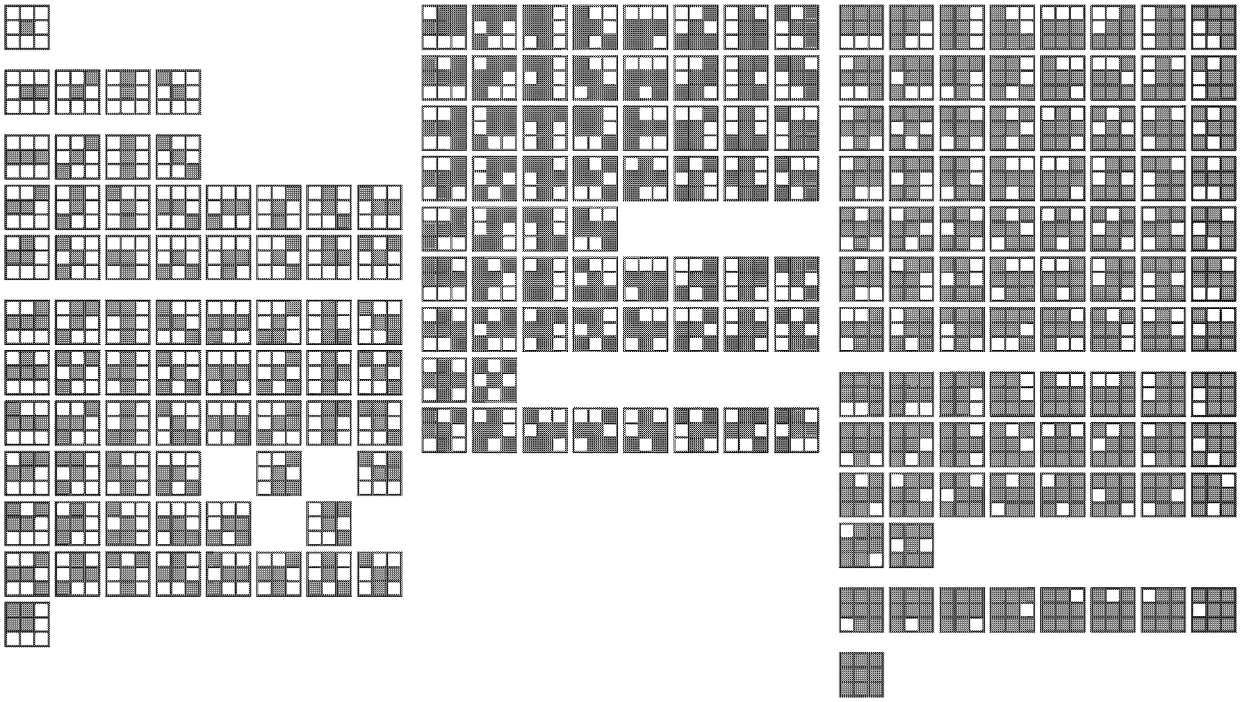

Figure 1: 223 mask patterns (3 × 3)

reference points. In this case, 2^8 patterns are created. By eliminating the patterns which are equivalent by the shift, the 223 mask patterns shown in Fig. 1 are obtained.

Feature values are calculated by scanning the image with the 223 mask patterns and by computing the sums of the products of the intensities of the corresponding pixels. Thereby, 223 features (a 223-dimensional feature vector) are extracted from one image. The features are obviously shift-invariant. Although we can compute the sum or difference of the intensities instead of the product, calculating the product showed the highest accuracy in a preliminary experiment.

The features extracted from 3 × 3 mask patterns consist of high frequency components of the image. The features of low frequency components can be obtained by enlarging the mask size. However, if all combinations of reference points are simply considered in a large region, the number of created mask patterns becomes huge, e.g., 2^{24} mask patterns in a 5 × 5 region. To avoid this situation, as shown in Fig. 2, the positions of the reference points are restricted to eight points that are located on the outer frame of masks at equal intervals. Mask patterns are created with combinations of these eight points. This approach allows the extraction of features of the low frequency components from the masks of various sizes without a deviation in the direction. It avoids an explosion in the number of patterns.

This study constructs multi-resolution features by concatenating the feature vectors that are extracted from each of three masks, 3 × 3, 5 × 5①, and 7 × 7①. Any number and combination of masks can be used. It is also possible to combine the masks ① and ② of the same size. However, the features extracted from the masks of different sizes comprise different frequency components. For that reason, they are considered to have low correlations and to be more useful for a recognition task. Therefore, this method uses masks of different sizes.

The reason that the mask① is adopted out of the same size masks ①, ② is that the patterns of the mask① were easily producible based on the 3 × 3 mask patterns, setting the positions of the reference points at intervals of one or two. In addition, a preliminary experiment did not indicate the difference in the discrimination power of the same-sized masks.

The feature values calculated by the proposed method represent a power spectrum of each mask pattern within the image. It is equivalent to expressing the base set of functions in the frequency analysis by the mask patterns. Compared with basis functions of the Fourier transform, the frequency and direction components correspond to the sizes and patterns of masks, respectively. The proposed method can analyze local regions in detail because it uses various two-dimensional patterns as well as directions.

The proposed method performs discrimination tasks using features described above along with "additional features" presented in 2.3.

2.2 Higher Order Local Autocorrelation Features

The N-th order autocorrelation functions, extensions of autocorrelation functions, are defined as

$$x(a_1, a_2, \cdots, a_N) = \int f(r)f(r + a_1) \cdots f(r + a_N)dr, \quad (1)$$

where $f(r)$ denotes the intensity at the observing pixel r, and $(a_1, a_2, \cdots, a_N)$ are Ndisplacements within a neighborhood. The original higher order local autocorrelations

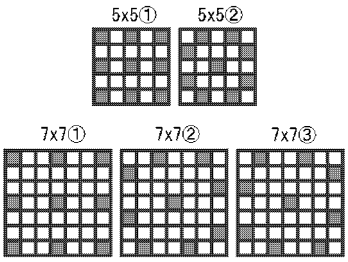

Figure 2: Positions of the reference points in masks of various sizes

that were used as image features were restricted to second-order and within a local 3 × 3 region[8]. They are represented as 25 mask patterns with 0, 1, and 2 reference points, which are identical to the first 25 patterns within the 223 mask patterns presented in Fig. 1. In terms of autocorrelation features, the 223 mask patterns are equivalent to the eighth-order autocorrelation features, where the reference positions and those numbers correspond to the displacements and the orders, respectively.

Utilization of the pyramidal images[4] and the multi-resolution approach using masks① shown in Fig. 2[3] were proposed as improvements of the original higher order local autocorrelation features. In another technique, which allows the duplication of displacement $(a_1, a_2, \cdots, a_N)$ in Eq. (1), the values were extracted from the same reference point of the mask pattern. All the autocorrelation features described above have been restricted to second-orders, mainly due to the high computational costs.

An approach which avoids explicitly computing the autocorrelations by exploiting the useful form of the inner products of autocorrelation functions has been proposed[9]. However, only two classes can be classified at one time by this technique because it uses a support vector machine. In [9], despite dealing with sixth-orders and a 9 × 9 region, the larger masks yielded lower accuracy rates: the best score was achieved with a 3 × 3 region. The main reason for this deterioration may be that the features became too numerous for the task because this technique created mask patterns with all the combination of the reference points in a large region. A huge number of features tend to include redundant and ineffective features, engendering lower recognition accuracy.

On the other hand, our proposed method prevents features from becoming too numerous: it restricts the number and positions of the reference points in a large mask.

2.3 "Additional Features"

In [3], features which have the same displacements within $(a_1, a_2, \cdots, a_N)$ in Eq. (1), were used for classification. Those features were calculated by applying the reference operations (product operations of the intensity) to the same pixels, including the central observing point of a mask pattern. The feature calculation formulae of masks with a 0 or 1 reference point are shown in Table 1, where $f(0)$ and $f(1)$ denote the intensities of the central observing point and the neighboring reference point, respectively.

Table 1: Feature calculation formulae of a mask pattern with a 0 or 1 reference points

	Number of reference points	
	0	1
	$f(0)$	$f(0) \times f(1)$
"Additional features"	$f(0) \times f(0)$	$f(0) \times f(0) \times f(1)$
"Additional features"	$f(0) \times f(0) \times f(0)$	$f(0) \times f(1) \times f(1)$

$f(0)$: Intensity of the central observing point
$f(1)$: Intensity of the neighboring reference point

Because these features were revealed to be useful in this study, we apply them for classification tasks as "additional features" in combination with the proposed features presented in 2.1. The number of reference operations is restricted to within twice to avoid rapid expansion of the number of "additional features" which result from the increase of operations. In all, 10 "additional features" are obtained from the first five mask patterns in Fig. 1 which have a 0 or 1 reference point (two features from each pattern). Therefore, 233 features (combined 223 features in 2.1 and 10 "additional features") can be extracted from each mask.

2.4 Feature Selection Method

Various features can be constructed by altering the combination of the mask patterns which are used for feature extraction. However, it is difficult to know the optimal features in advance because they depend on the task. It is also complicated and inefficient to search for the optimal features manually. For those reasons, we estimate the effective feature set based only on the given training samples. We apply a basic feature selection method. We use a total of 699 features, including "additional features," as the selection candidate features ("Original feature set") which are extracted from three masks, 3 × 3, 5 × 5①, 7 × 7①. The algorithm of feature selection from the "Original feature set" is shown below.

Step 1). Extract the "Original feature set" from each training sample image; then initialize the number of elements of the "Selected feature set" to 0.

Step 2). For each feature which has yet to be selected among the "Original feature set," calculate the class separability J of the training samples, when adding it to the existing "Selected feature set."

$$J = tr(\Sigma_W^{-1}\Sigma_B) \tag{2}$$

Σ_W: Within-class covariance matrices
Σ_B: Between-class covariance matrices

Step 3). Add the feature which maximizes the class separability calculated in Step 2 to the existing "Selected feature set."

Step 4). Repeat Step 2 and 3 until the [STOP POINT] mentioned below, and then, output the "Selected feature set."

Table 2: Details of experiment data

Test	Classes	Image size	Number of samples	
			Training	Test
00	24	128×128	240(24×10)	240(24×10)
01	24	64×64	1056(24×44)	1056(24×44)
14	68	128×128	680(68×10)	1360(68×10×2)

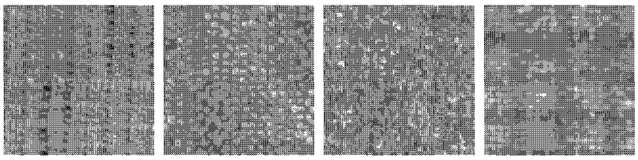

Figure 3: Examples of texture images

[**STOP POINT**] Classify the given training samples by M "Selected feature sets," which have elements of i to $i+M-1$ $(i = 1, 2, \cdots)$, and calculate the M recognition rates. Approximate a sequence of the M rates using a regression line; then compute its slope. Set the STOP POINT as that point just before the slope becomes zero.In this study, we set $M = 40$, based on a preliminary experiment.

3 Experiments

We performed texture classification experiments to assess the ability of the proposed features and the effectiveness of applying a feature selection method. We also examined the availability of the features in the real world through classification of the texture images that are illuminated with different light sources.

In these experiments, we used linear discriminant analysis for classification. Features were scaled so that the mean value of training samples was 1 in each dimension.

3.1 Image Data

We used the "Outex" texture database[6], which is a publicly available framework for experimental evaluation of texture analysis algorithms (http://www.outex.oulu.fi/outex.php). From Outex, we selected three test suites, Outex_TC_00000, 00001 and 00014 (Test00, Test01, and Test14). Table 2 shows details of those test suites. In each test suite, test samples were different from training samples. Examples of texture images are shown in Fig. 3.

Test00 and Test01 have 100 classification problems with the different set of training and test samples. Evaluations are performed by each average recognition rate. In Test14, to assess the robustness against illumination changes, three different illuminants were utilized. An illuminant "2856K incandescent CIE A" was used to capture training sample images; the others ("2300K horizon sunlight" and "4000K fluorescent TL84") were used for test sample images.

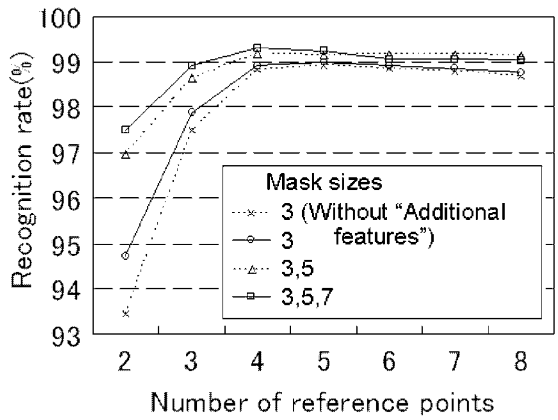

Figure 4: Classification by the multi-resolution features with a different number of reference points (Test01)

Table 3: Classification rates by the multi-resolution features with different number of reference points (Test01)

Mask sizes	2	3	4	5	6	7	8 [%]
$3^{(-)}$	93.4	97.5	98.8	98.9	98.9	98.8	98.7
3	94.7	97.9	98.9	99.0	98.9	98.8	98.8
3,5	97.0	98.6	99.2	99.1	99.2	99.2	99.1
3,5,7	97.5	98.9	***99.3***	99.2	99.0	99.1	99.0

$3^{(-)}$: Features without the "additional features"

All samples were grayscale images (original color images of Test14 were transformed). Each of them was normalized to have an average intensity of 128 and a standard deviation of 20.

3.2 Classification by The Proposed Features

Figure 4 and Table 3 show results of Test01 performed by the features which have resolutions of 1 to 3 and are constructed by the proposed masks with reference points of 2 to 8, including the "additional features." In these experiments, we set the number of reference points as equal between masks of different size that are used together.

Figure 4 illustrates the tendency for the recognition rate to rise concomitant with the increase in the number of reference points and the number of resolution levels used. The single resolution features extracted from the 3 × 3 mask patterns which have up to two reference points, without the "additional features," are equivalent to the original 25-dimensional autocorrelation features[8]. Their recognition rate became 93.4%. The result of their improved method[3, 4], which adds the "additional features" to the original features and constructs the multi-resolution features, was 97.5%.

The proposed features, which offer more reference points than conventional methods, improved the accuracy remarkably. Single resolution features of five reference points obtained 99.0%; the three-resolution features of four reference points achieved the best rate of 99.3%.

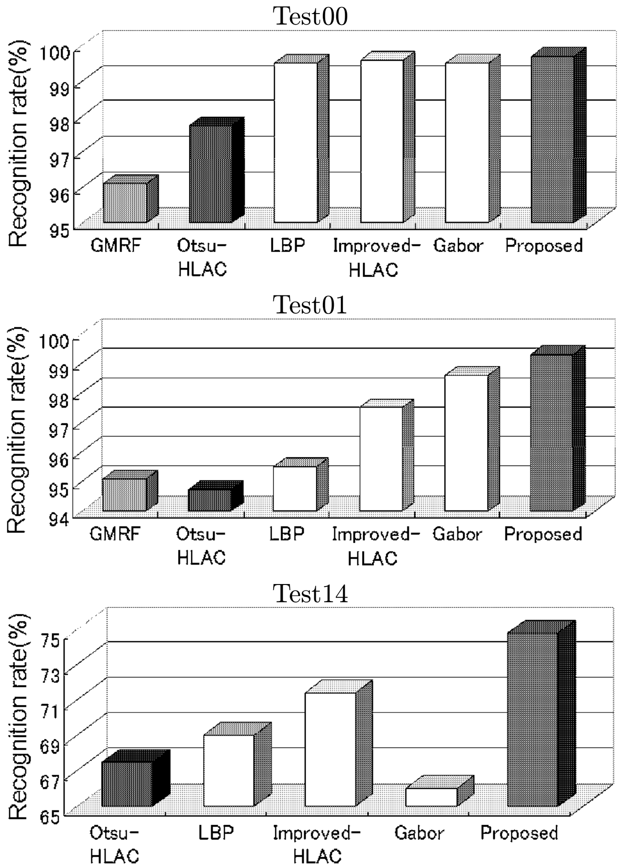

Figure 5: Comparison results of the proposed features with other features

Table 4 and Fig. 5 show results of Test00, Test01, and Test14 for several features: proposed features (Proposed), original higher order local autocorrelations (Original-HLAC), conventional improved features of them (Improved-HLAC), and other features (GMRF, LBP, Gabor) reported by [5, 6, 7]. Classifications were performed by linear discriminant analysis for Proposed, Original-HLAC and Improved-HLAC. The reported features adopted the 3-NN method (dissimilarity measures were log-likelihood for LBP and the Mahalanobis distance for Gabor.)

In all three tests, the proposed method achieved the best rates, 99.7% (99.6% with Improved-HLAC) in Test00, 99.3% (98.6% with Gabor) in Test01, 74.8% (71.4% with Improved-HLAC) in Test14. The high accuracy of the proposed features in Test14, where different illuminants were utilized, indicates the proposed method's feasibility for use in the real world.

The proposed method allows more detailed analysis than the conventional autocorrelation features because it uses a wider variety of mask patterns. Moreover, analysis of a larger region is achieved by this multi-resolution approach. These techniques raise the recognition accuracy and produce better results than other methods.

Table 4: Comparison results of the proposed features with other features

	Test00	Test01	Test14
Proposed	***99.7%***	***99.3%***	***74.8%***
Multi-resolutions	2	3	2
Reference points	3	4	8
Original-HLAC	97.2%	93.4%	64.1%
Improved-HLAC	99.6%	97.5%	71.4%
GMRF	96.1%	95.1%	N/A
LBP	99.5%	95.5%	69.0%
Gabor	99.5%	98.6%	66.0%

HLAC: Higher order local autocorrelations
GMRF: Gaussian Markov Random Field
LBP: Local Binary Pattern

3.3 Applying a Feature Selection Method

We assessed the ability of the features which are selected from the proposed features by the method mentioned in 2.4, using the same three tests as in 3.2. As described in 2.4, 699 features in all were used as the selection candidate features ("Original feature set"). In Test00 and Test01, we performed feature selection and classification test 100 times; then evaluated the average recognition rates.

Fig. 6 shows the transition of the recognition rates of traininig samples and test samples, when adding selected features one by one in the first problem of Test01. For reference, although the algorithm described in 2.4 halted feature selection at the point of the □ mark, this figure shows the whole transition until all 699 features have been added. The x-axis (horizontal) and y-axis (vertical) indicate the number of features used for classification and the obtained recognition rate, respectively.

As features are added, the recognition rate of the traininig samples continues to rise. On the other hand, the recognition rate of the test samples also rises at the beginning of feature addition, but from a certain point, it shows a declining tendency. This is considered to result from adding redundant and ineffective features.

Although it is appropriate to halt feature selection just before the recognition rate of the test samples falls, estimating the optimal stop point only from the training samples is difficult. In the example of Fig. 6, the feature selection stoped at the point of the □ mark, where the original 699 features were reduced to 157 and recognition rate rose to 99.1% from 98.3%.

The line with ● marks in Fig. 6 shows the recognition result of the features which were constructed without selection. Those features were three-resolution features and were extracted by the masks with 2 to 8 reference points. They also included the "additional features." Each ● mark corresponds to the features of 2 to 8 reference points. Table 5 shows the correspondence of the number of reference points and the number of features obtained.

As shown in Table 6 and Fig. 7, applying a feature selection method raised the recognition rates in all three tests. The number of features were reduced from original 699 to 70 in Test00, 154 in Test01, and 200 in Test14. These improvements are considered to be led by deleting the re-

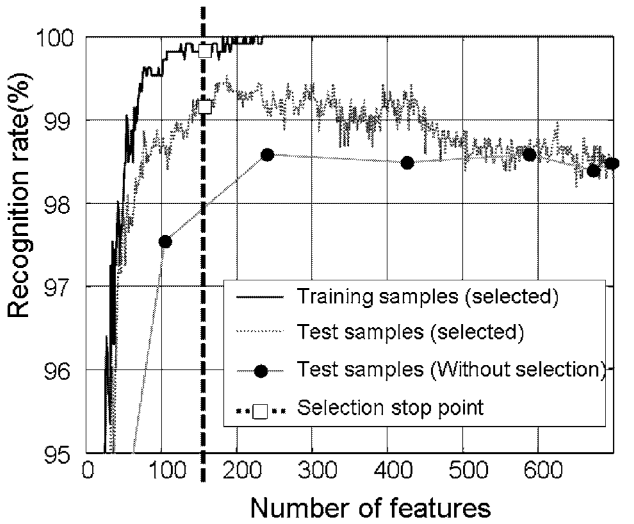

Figure 6: Transition of the recognition rate by applying the feature selection method (Test01)

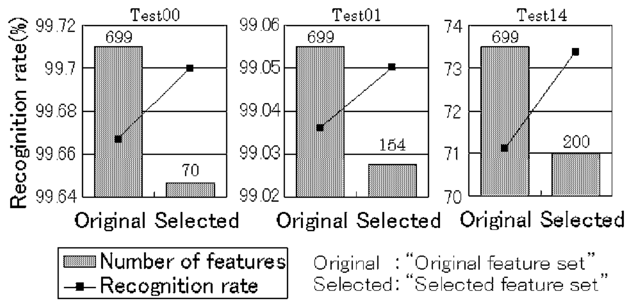

Figure 7: Results of applying the feature selection method

dundant features included in the "Original feature set," and using the effective features efficiently. It is also considered that reduction of the features avoids the "Hughes phenomenon,"[2], i.e., the classification accuracy declines as the number of features becomes excessive to the fixed number of samples.

The selected features differed with each problem. The numbers of those features also differed. This indicates that the appropriate feature selections were carried out for every problem. In Test14, although the features were selected under the different illuminant, classification task was executed well. Decreasing the number of features reduces the computational cost and memory required to calculate the feature values.

Table 5: Correspondence of the number of reference points and the number of three-resolution features

P	2	3	4	5	6	7	8
N	105	240	426	588	672	696	699

P: Number of reference points
N: Number of three-resolution features

Table 6: Results of applying the feature selection method

	Test00		Test01		Test14	
	rate%	N	rate%	N	rate%	N
Original	99.7	699	99.0	699	71.1	699
Selected	***99.7***	***70***	***99.1***	***154***	***73.4***	***200***

N: Number of features

4 Conclusion

We proposed a simple feature extraction method that is easy for hardware implementation. Mask patterns of 3 × 3 pixel size are represented by 223 patterns. They include second-order autocorrelation features. In the proposed method, mask patterns of several sizes which are based on the 3 × 3 mask are created. By combining those features, we construct multi-resolution features. In texture classification tests, the proposed features performed better than the second-order autocorrelations, Gabor, and other features. The proposed features also performed well in a test in which different illuminants are utilized to capture the images. These results indicate that the proposed method offers high practicality in the real world. Moreover, improvement of the recognition accuracy and considerable reduction of the number of features were realized through application of a feature selection method. Until now, the autocorrelation features have been applied to character recognition[8], facial discernment[1, 4], etc. The proposed method is also expected to demonstrate its high performance in various tasks.

References

[1] F. Goudail. et al., "Face recognition system using local autocorrelations and multiscale integration," IEEE Trans. PAMI, vol.18, No.10, pp.1024–1028, 1996.

[2] G. F. Hughes. On the mean accuracy of statistical pattern recognizers, IEEE Trans. Information Theory, vol.14, pp.55–63, 1968.

[3] O. Hasegawa. et al., Texture analsis using higher order local autocorrelation, The 55th National Conference of IPSJ, vol.2, pp.258–259, 1997, (in Japanese).

[4] T. Kurita. et al., A face recognition method using higher order local autocorrelation and multivariate analysis, Proc. ICPR, vol.2, pp.213–216, 1992.

[5] T. Mänpää. et al., Separating Color and Pattern Information for Color Texture Discrimination, Proc. ICPR, vol.1, pp.668–671, 2002.

[6] T. Ojala. et al., Outex-New framework for empirical evaluation of texture analysis algorithms, Proc. ICPR, vol.1, pp.701–706, 2002.

[7] T. Ojala. et al., Multiresolution gray-scale and rotation-invariant texture classification with Local Binary Patterns, IEEE Trans. PAMI, vol.24, No.7, pp.971–987, 2002.

[8] N. Otsu. Mathematical Studies on Feature Extraction in Pattern Recognition, Researches of the ETL, No.818, 1981, (in Japanese).

[9] V. Popovici. et al., Higher order autocorrelations for pattern classification, Proc. ICIP, pp.724–727, 2001.