Texture Recognition from Positions of the Theory of Active Perceptions

Vasiliy Gai(✉), Pavel Rodionov, Maxim Derbasov, Dmitriy Lyakhmanov, and Alla Koshurina

Nizhny Novgorod State Technical University N. A. R. E. Alekseev, K. Minin Street, 24, Nizhny Novgorod 603950, Russian Federation
vailiy.gai@gmail.com

Abstract. Recognition of textures is one of the topical tasks of computer vision. The key step in solving this problem is the formation of feature description of the texture image. A new approach to the formation of texture features based on the theory of active perception is proposed. The results of a computational experiment based on the Brodatz-32 database are presented, and the accuracy of the classification is demonstrated. The application of the proposed feature systems for recognition of snow and land textures in the solution of the problem of auto piloting in complex natural and climatic conditions is considered.

Keywords: Texture recognition · Theory of active perception · Auto piloting

1 Introduction

The task of recognizing textures is one of the fundamental problems in the field of computer vision and image processing. Texture recognition is used for automatic and automated analysis of medical images, object recognition, environmental modeling and image search in databases. The methods for recognizing textures also find their application in solving the problem of auto piloting in complex natural and climatic conditions.

The structure of the system for recognizing textured images can be represented as a set of three stages: image preprocessing, formation of image feature system and making decision.

The preliminary processing of the image usually consists in applying to the image of the filter suppressing noise. Often the implementation of this stage is not fulfilled. In this case, the responsibility for the noise invariance is shifted to the method of forming feature description.

When forming the feature description of a textured image, a wavelet transform is used, a method of analyzing independent components, Markov random fields and so on. In [1], it is proposed to use the Gabor filter bank to calculate the feature description. However, the filters included in the Gabor filter bank

B. Kryzhanovsky et al. (eds.), *Advances in Neural Computation, Machine Learning, and Cognitive Research*, Studies in Computational Intelligence 736,
DOI 10.1007/978-3-319-66604-4_15

are not orthogonal. The method of forming an image description based on local binary patterns is presented in [2]. It is based on the calculation of the sign of the difference between the brightness of neighboring samples. This method is not resistant to noise due to the use of threshold operation. The modification of the method of local binary patterns is known – the method of local ternary patterns [3]. The description formed on the basis of local ternary patterns is noise-resistant, but not resistant to brightness level changes, since the algorithm for generating the feature description uses fixed predetermined thresholds.

When solving the problem of classification of textured images on the basis of a well-known feature description, the method of K-nearest neighbors and artificial neural networks are often used.

This paper is devoted to solving the problem of recognizing textured images from the perspective of the theory of active perception (TAP).

2 Using TAP in Image Recognition

The basic transformation of TAP is a U-transformation, which is realized in two stages [4]. In the first stage, the Q-transformation is applied to the image, after which we obtain a matrix of visual masses m with a size of 4×4 elements. In the second stage, the set of filters $F = \{F_i\}$ is applied to the result of the Q-transformation.

The filter element can take the values "+1" (dark areas) and "−1" (light areas). Structurally, these filters are similar to the Walsh filters of the Harmuth system. The specificity of using these filters is that they are applied after the implementation of the Q-transformation.

In the TAP, to each filter F_i the binary operator V_i is put in correspondence. In this case the operator V_i or $\overline{V_i}$ corresponds to the component $\mu_i \neq 0$ of the vector μ_i depending on the sign of the component.

Defining the operations of set-theoretic multiplication and addition on the set $\{V_i\}$, an operation analogous to negation, two elements: $1 - V_0$ $0 - \overline{V_0}$ we obtain the algebra of signal description in Boolean functions: $AV =< \{V_i\} : +, \times >$. For any $V_i, V_j, V_k \in V$, the laws of commutativity, associativity, idempotency, and distributivity are fulfilled.

A group algebra is formed on the set of operators:

1. the family of algebraic structures P_n of cardinality 35, called complete groups, are formed on triples of operators (V_i, V_j, V_k);
2. the family of algebraic structures P_s of power 105 (for 16 filters), called closed groups, is formed on the four operators (V_i, V_j, V_p, V_m);

Comparing TAP with the known approaches to the formation of the feature description of the signal, we can note the following:

1. in comparison with the wavelet transform and the Fourier transform, TAP makes it possible to calculate, with respect to spectral coefficients, signs of a higher level (due to the use of group algebra);

2. in comparison with the models of deep learning in TAP, the feature description is calculated without using training, but by predefined templates;
3. only the addition and subtraction operations are used in calculating the U-transformation.

3 Formation of Feature Description of a Textured Image

The algorithm for forming the feature description of an image consists in combining descriptions of individual areas of the original image, obtained on the basis of complete or closed groups, into histograms of complete and closed groups. The algorithm for forming the feature description can be written as follows:

$$\begin{aligned}
&\forall i = 1 : s_h : (N - h)\\
&\quad \forall j = 1 : s_w : (M - w)\\
&\qquad I_s = I[i : (i + s_h - 1); j : (j + s_w - 1)];\\
&\qquad I_G = H[I_G, \Gamma];\\
&\qquad I_D[I_G] = I_D[I_G] + 1.
\end{aligned}$$

The following notations are used in the algorithm record: I_D – the image description obtained during the operation of the algorithm, I_s is the region of the image over which the I_G description is formed, the size of the region I_s – $h \times w$ samples, Γ is the type of the description being formed: P_{nm} – complete groups on the operation of multiplication, P_{na} – complete groups on the addition operation, P_s – closed groups. The I_G description can be obtained in the form of complete or closed groups. The value of the shift step of the area I_s in the

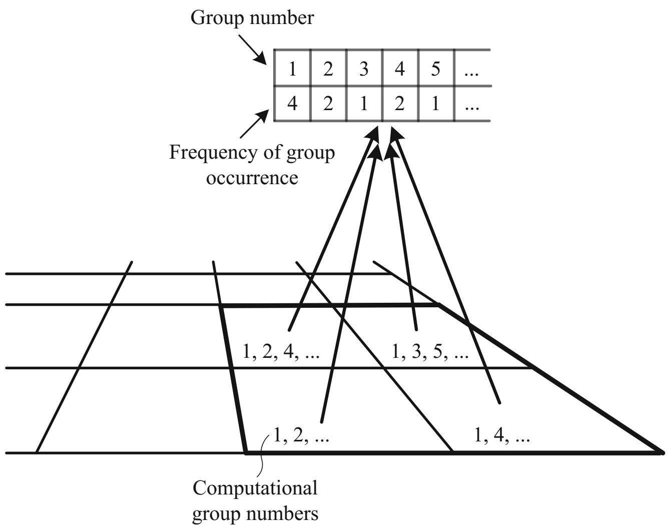

Fig. 1. The scheme of calculation of the histogram of groups

Table 1. Results of the use of histograms of complete and closed groups

Number of images in the sample	Feature descriptions used	Image area size I_s	Accuracy of classification (in %)	Time of feature description calculation (in ms)
$32 \times 16 = 512$	P_s	64×64	97	4
$32 \times 36 = 1024$	$P_s + P_{na}$	32×32	97	10
$32 \times 64 = 2048$	$P_s + P_{nm} + P_{na}$	32×32	94	5

image I is s_h of horizontal and s_w vertical counts of samples. The realization of the operator H for the calculation of complete and closed groups by the image is given in [4].

Figure 1 shows the scheme for calculating the group histogram.

4 Computational Experiment

As the initial data for the implementation of the computational experiment, the Brodatz-32 texture image database is used. When solving the classification problem, the support vector method with a linear kernel is used.

The results of estimating the accuracy of classification using various feature descriptions are given in Table 1. To form a sample, each texture image is divided into 16, 32 and 64 equal parts. The values of the parameters s_h and s_w are 1/4 of the size of the area.

The results of testing the proposed approach to the formation of the feature description to various distortions of the classified image are given in Tables 2 and 3. As the feature description, the description obtained after merging the histograms of complete (on addition and multiplication) and closed groups was used. This feature description provides the best results when classifying textured images without distortion.

The size of the classified image is 256×256 samples, the size of the sample is 512 images, the value of the window shift in the s_w (s_h) image is 32 samples. During the testing, as a test sample the distorted images were used, and as a training sample – the entire database of images of textures.

Table 2 shows the results of the evaluation of classification accuracy when the classified image is rotated by a given angle. Table 3 shows the results of experiments to study the invariance of the proposed feature description to the distortion of the classified image by normal noise.

Table 4 shows the results of the classification of the Brodatz-32 database on the basis of the known methods. Table 5 shows the evaluation of the noise invariance of the known systems of features (in %).

Table 2. Evaluation of the invariance to rotation

Angle of rotation (in degrees)	1	2	3	4	5	6	7	8	9	10	15	20
Accuracy of classification (in %)	94	93	93	89	93	91	90	89	87	87	70	58

Table 3. Evaluation of the noise invariance

Signal to noise ratio (in decibels)	20	15	10	5	0
Accuracy of classification (in %)	92	68	59	31	15

Table 4. Accuracy of classification based on the known methods

Reference	Feature description of the texture image	Algorithm of classification	Accuracy (in %)
[2]	Gabor filters	KNN	96
[2]	Local binary patterns	KNN	98
[3]	Convolutional neural network	Convolutional neural network	91
[5]	Local ternary patterns	SVM	95
[5]	Local high order statistics	SVM	99

Table 5. Invariance to distortion by noise of the known feature descriptions

Invariance to distortion by noise/Feature descriptions	20	10	5
SIFT	87	55	32
LBP	85	53	20

Comparing the proposed approach with the known ones, we can note the following:

1. the filters used in constructing the feature description of a textured image, in contrast to Gabor filters, are orthogonal;
2. when forming a feature description, in comparison with Local Binary Patterns and Local Ternary Patterns, there is no need to set predefined thresholds;
3. the invariance of the proposed feature descriptions to the distortion of the classified image by normal noise is comparable to the invariance to normal noise of LBP and SIFT features [2];
4. in [2] it is indicated that the average time for calculating LBP features by the image is 0.6 ms, based on Gabor filters – 197 ms (Intel Core i5-2400, C++); The time of calculation of the proposed systems of features depends on the size of the processed image, the size of the I_s region, and the magnitude of the

shift of the I_s region over the image, and ranges from 1 to 19 ms (Intel Core i7-4790K, C++); Thus, by the speed of computation the proposed feature descriptions are as good as the known ones.

5 Conclusion

The article considers an approach to the formation of the feature description of a textured image based on the theory of active perception. The practical application of this feature description is the segmentation of the image obtained from the video camera to solve the tasks of autopilot in the absence of a road network. Testing of the proposed approach to the formation of the feature description is performed using the method of reference vectors based on the Brodatz-32 textured image database. We obtained the result showing the effectiveness of the proposed systems of characteristics, their invariance to various distortions of the classified images.

Acknowledgements. The work was carried out at the NNSTU named after R. E. Alekseev, with the financial support of the Ministry of Education and Science of the Russian Federation under the agreement 14.577.21.0222 of 03.10.2016. Identification number of the project: RFMEFI57716X0222. Theme: "Creation of an experimental sample of an amphibious autonomous transport and technological complex with an intelligent control and navigation system for year-round exploration and drilling operations on the Arctic shelf."

References

1. Rosniza, R., Nursuriati, J.: Texture feature extraction using 2-D Gabor Filters. In: IEEE Symposium on Computer Applications and Industrial Electronics, ISCAIE 2012, pp. 173–178 (2012)
2. Chen, J., Kellokumpu, V., Zhao, G., Pietikinen, M.: RLBP: robust local binary pattern. In: Proceedings of the British Machine Vision Conference (2013)
3. Sharma, G., Hussain, S., Jurie, F.: Local higher-order statistics (LHS) for texture categorization and facial analysis. Lecture Notes in Computer Science, vol. 7578. Springer, Heidelberg (2012)
4. Utrobin, V.A.: Physical interpretation of the elements of image algebra. Phys. Usp. **47**, 1017–1032 (2004)
5. Hafemann, L.G.: An analysis of deep neural networks for texture classification. Dissertation presented as partial requisite to obtain the Masters degree, Curitbia, Brazil (2014)